

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 14:54:29 ; Search time 3778 Seconds
(without alignments)
12737.656 Million cell updates/sec

Title: US-09-955-732-1
Perfect score: 1980
Sequence: 1 atggccctgtgtacacagtgaag.....gagaggaggcgagcctga 1980

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estlin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_estl:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	em_gss_hum:*
18:	em_gss_inv:*
19:	em_gss_pln:*
20:	em_gss_vrt:*
21:	em_gss_fun:*
22:	em_gss_mam:*
23:	em_gss_mus:*
24:	em_gss_pro:*
25:	em_gss_rod:*
26:	em_gss_phg:*
27:	em_gss_vrl:*
28:	gb_gssl:*
29:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	804.6	40.6	1125	12	BM809903	BM809903 AGENCOURT
2	785	39.6	991	12	BM471207	BM471207 AGENCOURT
3	783.6	39.6	1061	12	BM906413	BM906413 AGENCOURT
4	767.2	38.7	875	10	BG752969	BG752969 602732427

5	760.6	38.4	938	13	BU190685	BU190685 AGENCOURT
6	758	38.3	812	10	BG750408	BG750408 602709243
7	717.4	36.2	917	13	BQ712114	BQ712114 AGENCOURT
8	711	35.9	750	10	BE531347	BE531347 601278540
9	706.4	35.7	947	13	BU537952	BU537952 AGENCOURT
10	706.2	35.7	790	10	BG326709	BG326709 602425578
11	698.6	35.3	846	10	BE563259	BE563259 601335565
12	674	34.0	729	12	BI086720	BI086720 602850066
13	673.8	34.0	726	10	BE869222	BE869222 601445030
14	670.8	33.9	952	13	BQ675874	BQ675874 AGENCOURT
15	664	33.5	910	10	BG169297	BG169297 602321027
16	653.6	33.0	932	10	BG340220	BG340220 602438408
17	652.4	32.9	1004	12	BM910493	BM910493 AGENCOURT
18	641.6	32.4	919	10	BE737087	BE737087 601304852
19	639.6	32.3	790	12	BM008005	BM008005 603617745
20	639.4	32.3	931	13	BQ898728	BQ898728 AGENCOURT
21	626.8	31.7	937	13	BQ718968	BQ718968 AGENCOURT
22	619.8	31.3	991	13	BQ707012	BQ707012 AGENCOURT
23	619.2	31.3	897	10	BE873337	BE873337 601450290
24	619	31.3	785	10	BE616235	BE616235 601279253
25	616.2	31.1	791	10	BE881145	BE881145 601492016
26	611.8	30.9	629	12	BM785529	BM785529 K-EST0063
27	596	30.1	1024	12	BM560192	BM560192 AGENCOURT
28	593.4	30.0	728	12	BI824278	BI824278 603040675
29	593	29.9	663	10	BG753072	BG753072 602732341
30	585.8	29.6	887	14	CA495391	CA495391 AGENCOURT
31	585.6	29.6	716	10	BE907514	BE907514 601497473
32	570.8	28.8	743	12	BM017302	BM017302 603643963
33	569.4	28.8	589	9	AL602806	AL602806 DKFZP686L
34	548.8	27.7	721	14	CB166643	CB166643 IBE603020
35	545.6	27.6	892	10	BF971565	BF971565 602239811
36	540.4	27.3	1228	12	BM921421	BM921421 AGENCOURT
37	540	27.3	672	12	BG831015	BG831015 602766736
38	537.6	27.2	669	10	BE563116	BE563116 601335734
39	537	27.1	678	12	BM047611	BM047611 603628949
40	535.4	27.0	943	12	BM563401	BM563401 AGENCOURT
41	528.4	26.7	531	10	BG751542	BG751542 602730640
42	524	26.5	890	10	BE901875	BE901875 601675476
43	510	25.8	622	10	BE898514	BE898514 601681423
44	497.8	25.1	1158	12	BM910866	BM910866 AGENCOURT
45	491	24.8	599	10	BE280203	BE280203 601158526

ALIGNMENTS

RESULT 1
BM809903 1125 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6581153 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5454822
DEFINITION 5', mRNA sequence.
ACCESSION BM809903
VERSION BM809903.1 GI:19126726
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1125)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LILNL at:
http://image.llnl.gov
Plate: LLCML1950 row: j column: 07
High quality sequence start: 12

Db 175 CTCCTGGGGCTGTCTCTGGGACTGCAGGATGGAGGGGACAATGATGATGCAGCAGAGGCC 234

QY 181 AGTTCTGAGCCACAAGAGAGGCCCCGAGTGAAGAGAGTCCACGGGGACCAAGACAGAC 240

Db 235 AGTTCTGAGCCACAAGAGAGGCCCCGAGTGAAGAGAGTCCACGGGGACCAAGACAGAC 294

QY 241 TTGGGGCAAGGATCCCAAGAGTCCCAAGAGCAAGAGAGAGAGGACAGCACTGCACCTC 300

Db 295 TTCGGGCAAGGATCCCAAGAGTCCCAAGAGCAAGAGAGAGAGGACAGCACTGCACCTC 354

QY 301 ATGTACAGCTGTCTGAGGCCCGCAGATGACATCCGCCCTGGCAGCCCAAGCTGAGGCAACC 360

Db 355 ATGTACAGCTGTCTGAGGCCCGCAGATGACATCCGCCCTGGCAGCCCAAGCTGAGGCAACC 414

QY 361 CGGCCTCCCCGGCTCCGCTACCTGCTGTAGTTTCTACAGAGAGAGAGAGTCTGAGC 420

Db 415 CGGCCTCCCCGGCTCCGCTACCTGCTGTAGTTTCTACAGAGAGAGAGAGTCTGAGC 474

QY 421 CAGGATGAGACGGTCTCTCTGGCGGTGAGTTTCCCTGACAGCAGCTCCCCAGCTGCACC 480

Db 475 CAGGATGAGACGGTCTCTCTGGCGGTGAGTTTCCCTGACAGCAGCTCCCCAGCTGCACC 534

QY 481 CTGGGCTGTCTTGGCCCTCTGAGTGAACACCAGCTGTACTTAGATGAGACGGGGGC 540

Db 535 CTGGGCTGTCTTGGCCCTCTGAGTGAACACCAGCTGTACTTAGATGAGACGGGGGC 594

QY 541 TTCAGCGTGACGTCTGTGGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 600

Db 595 TTCAGCGTGACGTCTGTGGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 654

QY 601 TGGGCCACACTCCAGGTATTGCAACCAAGCATGTGAGGAGCTCTAGGACAGCGGCT-TGT 659

Db 655 TGGGCCACACTCCAGGTATTGCAACCAAGCATGTGAGGAGCTCTAGGACAGCGGCTGTAC 714

QY 660 ACCGGGTGGCAGTGCCCTCACTGGGGCAGCCACTACAGAGAGAGACTGAACTCCGAACA 719

Db 715 CCGGGGTGGCAGTGCCCTCACTGGGGCAGCCACTACAGAGAGAGACTGAACTCCGAACA 774

QY 720 GAGCTGCCTCAATGATGATGACAGGCTATGGCCGACCTGAGTCTCTCGGCTCCAGCGC 779

Db 775 GAGCTGCCTCAATGATGATGACAGGCTATGGCCGACCTGAGTCTCTCGGCTCCAGCGC 834

QY 780 CGAGCCTGGCGGGTCTCTAC-AAACAGACAGATGAGACAGGCGATCCGTGCTGAGC-- 835

Db 835 CAAACCTGGCGGGTCTCTCCAAAACAGAACCGAAGGAGCAGCGGATCCCTGCTTAATTG 894

QY 836 TGTGGAAGTGTGTGATGTCACTGACCTGAGAGTGTCACTTCC--AAAGAGATCCGCC 892

Db 895 TGGAAAAAGGTTTGATGTCAAGGACCCGGAAAAAGGTCCTTCCAAAAAAACCCTCG 954

QY 893 AGGCTCTGAGCTGCGGCTGGGGGCTCCCC 922

Db 955 GGCTCTGGAACCTGCGGCTGGGGGCTTCCC 984

RESULT 3
BM906413
LOCUS
DEFINITION
AGENCY 6620017 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5590371
5', mRNA Sequence.
BM906413
BM906413
VERSION
KEYWORDS
EST.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1061)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg Ph.D.

Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12364 row: j column: 04
High quality sequence stop: 616.

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5590371"
/lab_host="DH10B"
/clone_lib="NH_MGC_125"
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;
Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool
of three ovaries, from females ranging in age from 38 to
49 yo. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 2.1 kb, insert size range 1-3.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 036."

```

BASE COUNT	238 a	344 c	302 g	177 t
ORIGIN				

Query Match	39.6%;	Score 783.6;	DB 12;	Length 1061;
Best Local Similarity	97.5%;	Pred. No. 1.2e-142;		
Matches 807;	Conservative	0;	Mismatches 19;	Indels 2;
				Gaps 1;

QY	1149	GGAGTCGGCCACAGCTGCTGCCGCACTGGAAAGAGACGCCACCGCTTCATTGAGGCTGCAAG	1208
Db	5	GGGATCGGCCACAGCTGCTGCCGCACTGGAAGAGACGCCACCGCTTCATTGAGGCTGCAAG	64
QY	1209	AGCACAGGGCACCCACAGTGTGCTGCTCCACTGCAAGATGGGGCTGACCGGCTCAGCGGCCAC	1268
Db	65	AGCACAGGGCACCCACAGTGTGCTGCTCCACTGCAAGATGGGGCTGACCGGCTCAGCGGCCAC	124
QY	1269	AGTGTGCGCTATGCCATGAAGCAGTACGAATGCAGCTTGAGCAGGCGCTGCGCCACGT	1328
Db	125	AGTGTGCGCTATGCCATGAAGCAGTACGAATGCAGCTTGAGCAGGCGCTGCGCCACGT	184
QY	1329	GCAGAGCTCCGGCCCATCGCCCGCCCAACCCCTGGCTTCTGCGCAGCTGCAGATCTA	1388
Db	185	GCAGAGCTCCGGCCCATCGCCCGCCCAACCCCTGGCTTCTGCGCAGCTGCAGATCTA	244
QY	1389	CCAGGGCATCTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCAGAAAGTGGTGGGT	1448
Db	245	CCAGGGCATCTCTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCAGAAAGTGGTGGGT	304
QY	1449	CTCCCCAGAGGAGCACCCACGCCCTGAAGTCTCTACACCATTCACACTCTTCCGCCAGA	1508
Db	305	CTCCCCAGAGGAGCACCCACGCCCTGAAGTCTCTACACCATTCACACTCTTCCGCCAGA	364
QY	1509	ACCTGAGGGTGTGGGGAGAGGAAGTTGTAGGCATGGAAGAGAGCCAGGACGCCCGAA	1568
Db	365	ACCTGAGGGTGTGGGGAGAGGAAGTTGTAGGCATGGAAGAGAGCCAGGACGCCCGAA	424
QY	1569	AGAAAGAGCCTGGGCCACGGGCCACGTATTAACCTCCGAGGGTCATGAGGTCCATCAGTCT	1628
Db	425	AGAAAGAGCCTGGGCCACGGGCCACGTATTAACCTCCGAGGGTCATGAGGTCCATCAGTCT	484
QY	1629	TCTGAGGCCCTCCTTGGAGCTGGAGAGCACTCAGAGACCAAGTACATGCCAGAGGTCTT	1688
Db	485	TCTGAGGCCCTCCTTGGAGCTGGAGAGCACTCAGAGACCAAGTACATGCCAGAGGTCTT	544
QY	1689	CTCTTCCACAGAGTCTTTCATGAAGAGCCTCTGCAGCCCTTCCACAGCTTGCAAGGAC	1748
Db	545	CTCTTCCACAGAGTCTTTCATGAAGAGCCTCTGCAGCCCTTCCACAGCTTGCAAGGAC	604

QY 1749 CAAGGGAGCCAGCAGGTGGACAGGGGGCTCAGCCTGCTGAGTCCCGCCAGTCACT 1808
| | | | |
Db 605 CAAGGAGGCCAGCAGGTGGACAGGGGGCTCAGCCTGCTGAGTCCCGCCAGTCACT 664
QY 1809 GGTATCCCTCCAGGGCAGTGCCTGTGTGGCCACCGGACCAGGCTTCCAGGAGCAGGA 1868
| | | | |
Db 665 GGTATCCCTCCAGGGCAGTGCCTGTGTGGCCACCGGACCAGGCTTCCAGGAGCAGGA 724
QY 1869 GCAGGGGCGAGGGGCGAGGGGCGAGAGAGCCCTGCATTCTCTACGCCAGGTTCCGG-- 1926
| | | | |
Db 725 GCAGGGGCGAGGGGCGAGGGGCGAGAGAGCCCTGCATTCTCTACGCCAGGTTCCGGGA 784
QY 1927 AAGGTGTGAGACAGAGCCAGCGTGCATGACAGTGAGAGAGGGCGAG 1974
| | | | |
Db 785 AGGTGTGAGACAGAGCCAGCGTGCATGACAGTGAGAGAGGGCGG 832

RESULT 4
BG752969 875 bp mRNA linear EST 15-MAY-2001
LOCUS 602732427F1 NIH_MGC_43 Homo sapiens CDNA clone IMAGE:4876014 5',
DEFINITION mRNA sequence.
ACCESSION BG752969
VERSION BG752969.1 GI:14063622
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 875)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: LLCM1758 row: e column: 07
High quality sequence stop: 818.
Location/Qualifiers
1. 875
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4876014"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 43"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT 184 a 267 c 275 g 149 t
ORIGIN

Query Match 38.7%; Score 767.2; DB 10; Length 875;
Best Local Similarity 97.0%; Pred. No. 1.7e-139;
Matches 846; Conservative 0; Mismatches 18; Indels 8; Gaps 6;

QY 196 GAGAGGCCCCGAGTGAGAGAGCTCCACGGGACACAGACAGACTTCGGGCAAGATCC 255
| | | | |
Db 1 GAGAGGCCCCGAGTGAGAGAGCTCCACGGGACACAGACAGACTTCGGGCAAGATCC 60

QY 256 CAGAGTCCCCAGAGAGAGAGAGAGAGAGACACCTGCACTCATGTACAGCTGCTG 315
| | | | |
Db 61 CAGAGTCCCCAGAGAGAGAGAGAGAGAGAGACACCTGCACTCATGTACAGCTGCTG 120
QY 316 AGGCCGAGATGATATCCCGCTGCGAGCCAGCTGAGAGCAACCCCGGCTCCCGGCTC 375
| | | | |
Db 121 AGGCCGAGATGATATCCCGCTGCGAGCCAGCTGAGAGCAACCCCGGCTCCCGGCTC 180
QY 376 CGCTACCTGCTGTAGTTTCTACACGAGAGAGAGAGAGTCTGAGCCAGATGAGACGCTC 435
| | | | |
Db 181 CGCTACCTGCTGTAGTTTCTACACGAGAGAGAGAGAGTCTGAGCCAGATGAGACGCTC 240
QY 436 CTCCTGGGCGTGAATTTCCCTGACACAGACTCCCCAGCTGCAACCTGGGCTGTCTTG 495
| | | | |
Db 241 CTCCTGGGCGTGAATTTCCCTGACACAGACTCCCCAGCTGCAACCTGGGCTGTCTTG 300
QY 496 CCCCTCTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGGCTTCAGCGTGAAGTCT 555
| | | | |
Db 301 CCCCTCTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGGCTTCAGCGTGAAGTCT 360
QY 556 GGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAG 615
| | | | |
Db 361 GGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAG 420
QY 616 GTATTGCACCAAGCATGTGAGGAGAGCTTAGGAGCGGCGCTGTACCGGGTGGAGTGCC 675
| | | | |
Db 421 GTATTGCACCAAGCATGTGAGGAGAGCTTAGGAGCGGCGCTGTACCGGGTGGAGTGCC 480
QY 676 CTCACCTGGGCGAGCCACTACACAGAGAGACTGAACCTCCGAACAGAGCTGCCTCAATGAG 735
| | | | |
Db 481 CTCACCTGGGCGAGCCACTACACAGAGAGACTGAACCTCCGAACAGAGCTGCCTCAATGAG 540
QY 736 TGGACGGCTATGGCCGACCTGAGTCTTGTGCGGCTCCAGCGCCGAGCGCTGGCGGTCC 795
| | | | |
Db 541 TGGACGGCTATGGCCGACCTGAGTCTTGTGCGGCTCCAGCGCCGAGCGCTGGCGGTCC 600
QY 796 TCAGAACAGAGCAGATGAGAGAGCGGATCCGCTGAGCTGTGAAAGTGTGATGTC 855
| | | | |
Db 601 TCAGAACAGAGCAGATGAGAGAGCGGATCCGCTGAGCTGTGAAAGTGTGATGTC 660
QY 856 AGTGACCTGGA-GAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTGAGCTGGCGCTGGG 914
| | | | |
Db 661 AGTGACCTGGAAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTGAGCTGGCGCTGGG 720
QY 915 GCTCCCCC-TCACAGATACCGTACTTCATGCACAAC--AGATGCTGCTGTGTGG 970
| | | | |
Db 721 GCTCCCCC-TCACAGATACCGTACTTCATGCACAACAGATGTGCTGTGTGTGG 780
QY 971 CACAGCGGAGCCGAGCCTCCCGCATCTTCCCCCACTTACCTGGGCTCAGAGTGAACG 1030
| | | | |
Db 781 CACAGCGGAGCCGAGCCTCCCGCATCTTCCCCCACTTACCTGGGCTCAGAGTGAACG 838
QY 1031 CAGCAAACTTGAGAGAGCTGCAGAGGAACAGG 1062
| | | | |
Db 839 CAG-AAACCTTGAGAGAGCTGCAGAGGAACAGG 869

RESULT 5
BUI90685 938 bp mRNA linear EST 04-SEP-2002
LOCUS AGENCOURT 7953319 NIH_MGC_39 Homo sapiens CDNA clone IMAGE:6107345
DEFINITION 5', mRNA sequence.
ACCESSION BUI90685
VERSION BUI90685.1 GI:22704669
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LCM2349 row: n column: 18
High quality sequence stop: 605.
Location/Qualifiers

FEATURES

source

1.938
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6107345"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_39"
/note="Organ: pancreas; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 207 a 306 c 276 g 149 t
ORIGIN

Query Match 38.4%; Score 760.6; DB 13; Length 938;
Best Local Similarity 97.3%; Pred. No. 3.4e-138;
Matches 795; Conservative 0; Mismatches 19; Indels 3; Gaps 2;

QY 1151 AGTCGGCCAGCTGTGCGCCACTGGAAGAGAGACGACCGCTTCATTGAGGCTGCAAG 1210
Db 1 AGTCGGCCAGCTGTGCGCCACTGGAAGAGAGACGACCGCTTCATTGAGGCTGCAAG 60
QY 1211 CACAGGGCACCCAGCTGTGCTCCACTGCAAGATGGCGCTCAGCCGCTCAGCGCCACAG 1270
Db 61 CACAGGGCACCCAGCTGTGCTCCACTGCAAGATGGCGCTCAGCCGCTCAGCGCCACAG 120
QY 1271 TGCTGGCCTATGCCATGACAGTACGATGACGCTTGAGAGAGGCCCTGCGCCACGTGC 1330
Db 121 TGCTGGCCTATGCCATGACAGTACGATGACGCTTGAGAGAGGCCCTGCGCCACGTGC 180
QY 1331 AGGAGCTCCGGCCCATCGCCGCCCAACCCCTGCTTCTGCGCCAGCTGCAGATCTACC 1390
Db 181 AGGAGCTCCGGCCCATCGCCGCCCAACCCCTGCTTCTGCGCCAGCTGCAGATCTACC 240
QY 1391 AGGGCATCTGACGGGCCAGCGCCAGAGCCATGCTGGGAGCAGAAAGTGGGTGCT 1450
Db 241 AGGGCATCTGACGGGCCAGCGCCAGAGCCATGCTGGGAGCAGAAAGTGGGTGCT 300
QY 1451 CCCCAAGAGAGACCCAGCCCTGAAGTCTTACACCAATCCCACTCTTCCGCCAGAAC 1510
Db 301 CCCCAAGAGAGACCCAGCCCTGAAGTCTTACACCAATCCCACTCTTCCGCCAGAAC 360
QY 1511 CTGAGGGTGTGGGAGAGAGAGTGTAGGCGATGAGAGAGAGCCAGGAGCCCGGAAAG 1570
Db 361 CTGAGGGTGTGGGAGAGAGAGTGTAGGCGATGAGAGAGAGCCAGGAGCCCGGAAAG 420
QY 1571 AAGAGCTGGGCCACGCGCACGTATAAACCCTCCGAGGGGTGATGAGTCCATCAGTCTTC 1630
Db 421 AAGAGCTGGGCCACGCGCACGTATAAACCCTCCGAGGGGTGATGAGTCCATCAGTCTTC 480
QY 1631 TGGAGCCCTCTTGGAGCTGGAAGACCTCAGAGACCAAGTACATGCGAGAGGTCTTCT 1690
Db 481 TGGAGCCCTCTTGGAGCTGGAAGACCTCAGAGACCAAGTACATGCGAGAGGTCTTCT 540
QY 1691 CTCCACAGAGTCTTACATGAAGAGCCTCTGAGCCCTTCCACAGCTTGCAAGAGACCA 1750

Db 541 CTTCGCCACGAGTCTTCAATGAAGAGCCTCTGACGCCCTTCCACAGACTTGCAAGACCA 600

QY 1751 AGGGAGGCCAGCAGGTGAGACAGGGGGCCCTCAGCCTGCCCTGAAGTCCGCCAGTCACTGG 1810

Db 601 AGGGAGGCCAGCAGGTGAGACAGGGGGCCCTCAGCCTGCCCTGAAGTCCGCCAGTCACTGG 660

QY 1811 TTACCTCCAGAGGAGTCCGCTGTGTGGCCAAACCGAACCCAGGCTTCCAGAGAGCAGAGC 1870

Db 661 TTACCTCCAGAGGAGTCCGCTGTGTGGCCAAACCGAACCCAGGCTTCCAGAGAGCAGAGC 720

QY 1871 AGGGAGAGGGGCGAGGGGC--AGGAGAGCCCTGCATTTCTCTACGCCAGGTT-CCGGA 1927

Db 721 AGGGAGAGGGGCGAGGGGCAGGGGCAAGAAACCCCTGCATTTCTCTACGCCAGGTTCCCGGA 780

QY 1928 AGGTGTGAGACAGAGGCCACGCTGCATGACAGTGGAGA 1964

Db 781 AAGTGTGAGACAGAGGCCCGCTGCATGGAATGGAAA 817

RESULT 6

BG750408 812 bp mRNA linear EST 15-MAY-2001
LOCUS 602709243F1 NIH_MGC_43 Homo sapiens CDNA clone IMAGE:4845707 5',
DEFINITION mRNA sequence.

ACCESSION BG750408
VERSION BG750408.1 GI:14061061

KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

AUTHORS 1 (bases 1 to 812)
TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished

CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

Plate: LCM1685 row: f column: 12
High quality sequence stop: 792.
Location/Qualifiers

FEATURES

source

1.812
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4845707"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. |"

BASE COUNT 167 a 250 c 254 g 141 t
ORIGIN

Query Match 38.3%; Score 758; DB 10; Length 812;
Best Local Similarity 97.3%; Pred. No. 1.1e-137;
Matches 792; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 196 GAGAAGCCCCGAGTGAAGAGAGAGTCCACGGGACACAGACACTTGGGCAAGATCC 255
Db 1 GAGAAGCCCCGAGTGAAGAGAGAGTCCACGGGACACAGACACTTGGGCAAGATCC 60

QY	256	CAGAGTCCCCAGAAACGAGGAGGACAGAGGACAGCACTGCACTTCATGTGTACAGCTGCTG	315
Db	61	CAGAGTCCCCAGAAACGAGGAGGACAGAGGACAGCACTGCACTTCATGTGTACAGCTGCTG	120
QY	316	AGGCCGCAAGGATGACATCCGCCCTGGCAGCCCCAGCTGGAGGGCACCCCGGCTCCCGGGCTC	375
Db	121	AGGCCGCAAGGATGACATCCGCCCTGGCAGCCCCAGCTGGAGGGCACCCCGGCTCCCGGGCTC	180
QY	376	CGCTACCTGCTGTAGTCTTCTACACGAGAAGGAGAAGGCTTGAGCCAGATGAGACGCTC	435
Db	181	CGCTACCTGCTGTAGTCTTCTACACGAGAAGGAGAAGGCTTGAGCCAGATGAGACGCTC	240
QY	436	CTCCTGGGCGTGGATTTCCCTGACAGACGCTCCCCAGCTGCACCCTGGGCTGTCTTG	495
Db	241	CTCCTGGGCGTGGATTTCCCTGACAGACGCTCCCCAGCTGCACCCTGGGCTGTCTTG	300
QY	496	CCCCCTGTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGGCTTCAGCGTGAAGTCT	555
Db	301	CCCCCTGTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGGCTTCAGCGTGAAGTCT	360
QY	556	GGTGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACATGTGGCCACACTCCAG	615
Db	361	GGTGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACATGTGGCCACACTCCAG	420
QY	616	GTATTGCACCAAGCATGTGAGGACGCTCTAGGACGGGCTTTGACCGGATGGCAGTGCC	675
Db	421	GTATTGCACCAAGCATGTGAGGACGCTCTAGGACGGGCTTTGACCGGATGGCAGTGCC	480
QY	676	CTCACCCTGGGCCAGCCACTACCAGAGAGACTGAACTCCGAACAGAGCTGCTCAATGAG	735
Db	481	CTCACCCTGGGCCAGCCACTACCAGAGAGACTGAACTCCGAACAGAGCTGCTCAATGAG	540
QY	736	TGGACGGCTATGGCCCGACCTGGAGTCTCTGGGCTCCCAAGCCCGAGCCTGGCCGGTCC	795
Db	541	TGGACGGCTATGGCCCGACCTGGAGTCTCTGGGCTCCCAAGCCCGAGCCTGGCCGGTCC	600
QY	796	TCAGAACAGAGCAGATGAGAGCAGCGATCCGTGCTGAGCTGTGAAAGTGTGGATGTC	855
Db	601	TCAGAACAGAGCAGATGAGAGCAGCGATCCGTGCTGAGCTGTGAAAGTGTGATGTC	660
QY	856	AGTGACCTGAGAGTGTCACTTCCAAAGAGATCCGCCAAGGCTCTGGAAGTCCCGCTGGG	915
Db	661	AGTGACCTGAGAGTGTCACTTCCAAAGAGATCCGCCAAGGCTCTGGAAGTCCCGCTGGG	720
QY	916	CTCCCCCTCCAGCAGTACCCTGACTTCATGACAAACCAAGATGCTGCTGCTGCTGACAG	975
Db	721	TGTCCCCCTCCAGCAGTACCCTGACTTCATGACAAACCAAGATGCTGCTGCTGCTGACAG	779
QY	976	CGGACCGAGCCTCCCGCATCTTCCCCCACTCT 1009	
Db	780	C-GGAACGAGCCTCCCGATTTTCCCCCACTCTACT 812	

RESULT 7
 BQ712114
 LOCUS
 DEFINITION
 BQ712114 917 bp mRNA linear EST 16-JUL-2002
 AGENCOURT 8418352 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6281726
 5', mRNA sequence.
 BQ712114
 BQ712114.1 GI:21851013
 EST.
 Homo sapiens (human)
 Homo sapiens
 ORGANISM
 SOURCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 917)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Mark Watson

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2474 row: P column: 15
High quality sequence stop: 584.
location/Qualifiers
1..917

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6281726"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_113"
/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

```

BASE COUNT	196 a	-315 c	263 g	143 t
ORIGIN				
Query Match		36.2%	Score 717.4;	DB 13;
Best Local Similarity		99.2%	Pred. NO. 9e 130;	Length 917;
Matches 721; Conservative		0;	Mismatches 6;	Indels 0;
				Gaps 0;

QY	1254	CCGCTCAGCGGGCCACAGTGTCTGGCCTATGCCATGAAGCAGTACGAATGCAGCCTGGAGCA	1313
Db	1	CCGCTCAGCGGGCCACAGTGTCTGGCCTATGCCATGAAGCAGTACGAATGCAGCCTGGAGCA	60
QY	1314	GGCCCTGCGCCACGTGACGAGAGCTCCGGCCCATGCCCCGCCAACCTGGCTTCTGCG	1373
Db	61	GGCCCTGCGCCACGTGACGAGAGCTCCGGCCCATGCCCCGCCAACCTGGCTTCTGCG	120
QY	1374	CCAGCTGCAGATCTACCAAGGCGCATCCTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCA	1433
Db	121	CCAGCTGCAGATCTACCAAGGCGCATCCTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCA	180
QY	1434	GAAAGTGGGTGGGTCTCCACAGAGAGCACCCAGCCCTGAAGTCTCTACACCATTTCC	1493
Db	181	GAAAGTGGGTGGGTCTCCACAGAGAGCACCCAGCCCTGAAGTCTCTACACCATTTCC	240
QY	1494	ACCTCTTCCGCCCAAGACCTGAGGGTGTGGGGAGAGAGAAGTTTATGCGATGGAAGAGAG	1553
Db	241	ACCTCTTCCGCCCAAGACCTGAGGGTGTGGGGAGAGAGAAGTTTATGCGATGGAAGAGAG	300
QY	1554	CCAGGCAGCCCCGAAAGAAGAGCCTGGGCCACGGCCACGTATATAACCTCCGAGGGGTCA	1613
Db	301	CCAGGCAGCCCCGAAAGAAGAGCCTGGGCCACGGCCACGTATATAACCTCCGAGGGGTCA	360
QY	1614	GAGGTCCATCAGTCTTCTTGAGAGCCCTCTTGAGAGCTGAGAGACACTCAGAGACCACTGA	1673
Db	361	GAGGTCCATCAGTCTTCTTGAGAGCCCTCTTGAGAGCTGAGAGACACTCAGAGACCACTGA	420
QY	1674	CATGCCAGAGGTCTTCTCTTCCCAAGTCTTCAATGAAGAGCCTCGAGCCCTTCCC	1733
Db	421	CATGCCAGAGGTCTTCTCTTCCCAAGTCTTCAATGAAGAGCCTCGAGCCCTTCCC	480
QY	1734	ACAGCTTGCAAGACCAAGGGAGGCCACAGGTGACACAGGGGCTCAGCCTGCCCTGAA	1793
Db	481	ACAGCTTGCAAGACCAAGGGAGGCCACAGGTGACACAGGGGCTCAGCCTGCCCTGAA	540
QY	1794	GTCCCGCCAGTCAGTGTTACCTTCCAGGGCAGTGCCTGTGTGCCCAACCGGACCCAGGC	1853
Db	541	GTCCCGCCAGTCAGTGTTACCTTCCAGGGCAGTGCCTGTGTGCCCAACCGGACCCAGGC	600
QY	1854	CTTCCAGGAGCAGAGCAGGGGCGAGGGGCAAGGGGCAAGGGGCAAGCCCTGCATTTCTCTAC	1913

Db 601 CTTCAGAGACGACGAGGCGCAGGCGCAGGAGAGCCCTGCATTCTCTAC 660

QY 1914 GCCCAGGTTCCGGAAGTGTGTGAGACAGGCGCAGCGTGCATGACGTGAGAGAGGCGCA 1973

Db 661 GCCCAGGTTCCGGAAGTGTGTGAGACAGGCGCAGCGTGCATGACGTGAGAGAGGCGG 720

QY 1974 GGCCTGA 1980

Db 721 AGCCTGA 727

RESULT 8
BE531347 750 bp mRNA lineat EST 09-AUG-2000
LOCUS 601278540F1 NIH_MGC_39 Homo sapiens CDNA clone IMAGE:3610616 5',
DEFINITION mRNA sequence.
BE531347
ACCESSION BE531347.1 GI:9759906
VERSION BE531347.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 750)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM266 row: P column: 09
High quality sequence start: 3
High quality sequence stop: 750.
FEATURES
source location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3610616"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_39"
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."

BASE COUNT 158 a 240 c 232 g 120 t

ORIGIN

Query Match 35.9%; Score 711; DB 10; Length 750;
Best Local Similarity 100.0%; Pred. No. 1.5e-128;
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1100 ACTTCTACCTGAGCGCTTACCTACCAATGTGCGCCTCTGGAGTAGAGAGTGGCCC 1159

Db 40 ACTTCTACCTGAGCGCTTACCTACCAATGTGCGCCTCTGGAGTAGAGAGTGGCCC 99

QY 1160 AGCTGCTGCGCACTGGAAGAGAGCAGCGCTTCAATGAGGCTGCAAGAGCAGAGGCA 1219

Db 100 AGCTGCTGCGCACTGGAAGAGAGCAGCGCTTCAATGAGGCTGCAAGAGCAGAGGCA 159

QY 1220 CCCACGTGCTGCTCACTGCAAGATGGGCGTCAAGCGGCGCCACAGTGTCTGGCCT 1279

Db 160 CCCACGTGCTGCTCACTGCAAGATGGGCGTCAAGCGGCGCCACAGTGTCTGGCCT 219

QY 1280 ATGCCATGAAGCAGTACCAATGCAGCCTGAGCAGAGCCCTGCGCCACGTGACAGAGCTCC 1339

Db 220 ATGCCATGAAGCAGTACCAATGCAGCCTGAGCAGAGCCCTGCGCCACGTGACAGAGCTCC 279

QY 1340 GGGCCATGCGCCCGCCCAACCTGCTTCTCTGCGCCAGCTGCAAGATCTACCAAGGCAATCC 1399

Db 280 GGGCCATGCGCCCGCCCAACCTGCTTCTCTGCGCCAGCTGCAAGATCTACCAAGGCAATCC 339

QY 1400 TGACGGCCAGCCCGCCAGAGCCATGTCTGGAGCAGAAAGTGGGTCTTCCCAAGAG 1459

Db 340 TGACGGCCAGCCCGCCAGAGCCATGTCTGGAGCAGAAAGTGGGTCTTCCCAAGAG 399

QY 1460 AGCACCAGCCCTGAAGTCTCTACACCATTCACCTCTTCCGCGCAGAACTGAGGGTG 1519

Db 400 AGCACCAGCCCTGAAGTCTCTACACCATTCACCTCTTCCGCGCAGAACTGAGGGTG 459

QY 1520 GTGGGAGAGAGAGTGTGTAGCATGGAAGAGAGCCAGCCGCGGAAAGAGAGCCTG 1579

Db 460 GTGGGAGAGAGAGTGTGTAGCATGGAAGAGAGCCAGCCGCGGAAAGAGAGCCTG 519

QY 1580 GGGCAGCGCCACGTATAAACCCTCCAGGGGTGTCATGAGTCCATCAGTCTTCTGAGCCCT 1639

Db 520 GGGCAGCGCCACGTATAAACCCTCCAGGGGTGTCATGAGTCCATCAGTCTTCTGAGCCCT 579

QY 1640 CCTTGAGCTGAGAGACACCTAGAGACCAGTGCATGCCAGAGGTCTTCTTCCACAG 1699

Db 580 CCTTGAGCTGAGAGACACCTAGAGACCAGTGCATGCCAGAGGTCTTCTTCCACAG 639

QY 1700 AGTCTTACATGAAGAGCCTCTGACGCCCTTCCACAGCCTTGCAAGAGCAGGAGGCC 1759

Db 640 AGTCTTACATGAAGAGCCTCTGACGCCCTTCCACAGCCTTGCAAGAGCAGGAGGCC 699

QY 1760 AGCAGGTGAGCAGGGGCGCTCAGCCTGCCCTGAAGTCCCGCCAGTCACTGG 1810

Db 700 AGCAGGTGAGCAGGGGCGCTCAGCCTGCCCTGAAGTCCCGCCAGTCACTGG 750

RESULT 9
BU537952 947 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT 10186579 NIH_MGC_107 Homo sapiens CDNA clone
DEFINITION IMAGE:6568457 5', mRNA sequence.
BU537952
ACCESSION BU537952.1 GI:22848393
VERSION BU537952
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 947)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2754 row: C column: 17
High quality sequence stop: 614.
FEATURES
source location/Qualifiers
1..947
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6568457"
/tissue_type="adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_107"
/note="Organ: breast; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC library."

BASE COUNT 204 a 314 c 271 g 153 t 5 others
ORIGIN

Query Match 35.7%; Score 706.4; DB 13; Length 947;
Best Local Similarity 99.3%; Pred. No. 1.3e-127;
Matches 740; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 1239 CAAGATGGGCGTCAGCCGCTCAGCGCCACAGTGTGCGCTATGCCATGAAGCAGTACGA 1298
DB 1 CAAGATGGGCGTCAGCCGCTCAGCGCCACAGTGTGCGCTATGCCATGAAGCAGTACGA 60
QY 1299 ATGCAGCCTGAGAGAGGCGCTGCGCCACGTCAGAGAGCTCCGGGCCATCGCCGCCCA 1358
DB 61 ATGCAGCCTGAGAGAGGCGCTGCGCCACGTCAGAGAGCTCCGGGCCATCGCCGCCCA 120
QY 1359 CCCTGGCTTCCTGCCACAGCTGAGATCTACAGGGCATCTTGACGGCCAGCCGACAGAG 1418
DB 121 CCCTGGCTTCCTGCCACAGCTGAGATCTACAGGGCATCTTGACGGCCAGCCGACAGAG 180
QY 1419 CCATGTCTGGAGAGAGAAAGTGGGTGGGTCTCCCCAGAGAGACACCCAGCCCTGAAGT 1478
DB 181 CCATGTCTGGAGAGAGAAAGTGGGTGGGTCTCCCCAGAGAGACACCCAGCCCTGAAGT 240
QY 1479 CTCTACACCATTTCCCACTCTTCCGCCAGAACTGAGGGTGTGGGAGAGAGAGTTGT 1538
DB 241 CTCTACACCATTTCCCACTCTTCCGCCAGAACTGAGGGTGTGGGAGAGAGAGTTGT 300
QY 1539 AGGCATGGAAGAGAGCCAGGAGCCCGGAAAGAGAGAGCTGGGCCACGGCCAGCTATAA 1598
DB 301 AGGCATGGAAGAGAGCCAGGAGCCCGGAAAGAGAGAGCTGGGCCACGGCCAGCTATAA 360
QY 1599 CCTCCGAGGGGTCATGAGTCCATCATCTTCTGAGCCCTCTTGAGCTGAGAGAGCAC 1658
DB 361 CCTCCGAGGGGTCATGAGTCCATCATCTTCTGAGCCCTCTTGAGCTGAGAGAGCAC 420
QY 1659 CTCAGAGACCAAGTACATGCCAGAGTCTTCTTCCAGAGTCTTCAATGAAGAGCC 1718
DB 421 CTCAGAGACCAAGTACATGCCAGAGTCTTCTTCCAGAGTCTTCAATGAAGAGCC 480
QY 1719 TCTGAGACCCCTTCCACAGCTTGCAAGAACCAAGGAGGCCAGAGTGAAGGGGGCC 1778
DB 481 TCTGAGACCCCTTCCACAGCTTGCAAGAACCAAGGAGGCCAGAGTGAAGGGGGCC 540
QY 1779 TCAGCCTGCGCTGAAGTCCCGCAGTCAAGTGTACCTCCAGGGCAGTGCCTGTGGC 1838
DB 541 TCAGCCTGCGCTGAAGTCCCGCAGTCAAGTGTACCTCCAGGGCAGTGCCTGTGGC 600
QY 1839 CAACCGGAGCCAGGCGCTTCCAGAGCAGAGAGAGGGGAGGGGAGGGGAGAGAGCC 1898
DB 601 CAACCGGAGCCAGGCGCTTCCAGAGCAGAGAGAGGGGAGGGGAGGGGAGAGAGCC 660
QY 1899 CTGCATTTCTCTACGCCCAAGTT-CCGGAAGGTGTGAGACA-GGCCAGCGTGCATGAC 1956
DB 661 CTGCATTTCTCTACGCCCAAGTTCCCGGAAGGTGTGAGACAAGGGCCAGCGTGCATGAC 720
QY 1957 AGTGAG-AGAGAGGGGAGGGCTGA 1980
DB 721 AGTGAGAGAGAGGGGAGGGCTGA 745

RESULT 10
BG326709 790 bp mRNA linear EST 27-FEB-2001
LOCUS BG326709

DEFINITION 602425578F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4563595 5',
mRNA sequence.
ACCESSION BG326709
VERSION BG326709.1 GI:13133146
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNCM1277 row: k column: 20
High quality sequence stop: 737.
Location/Qualifiers
1. 790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4563595"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_14"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

FEATURES

source

BASE COUNT 162 a 237 c 268 g 123 t
ORIGIN

Query Match 35.7%; Score 706.2; DB 10; Length 790;
Best Local Similarity 98.1%; Pred. No. 1.3e-127;
Matches 725; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 1 ATGCCCCGTGTCACAGTACGCGCTTCCGCCGCGGAGCGGCGCTCCAGCGCGTGGG 60
DB 50 ATGCCCCGTGTCACAGTACGCGCTTCCGCCGCGGAGCGGCGCTCCAGCGCGTGGG 109
QY 61 CCCTGGAGCAGCGCGTCCAGAGAGAGTGCATCCAGCGAAGGACAGAGCTTGGCGTG 120
DB 110 CCCTGGAGCAGCGCGTCCAGAGAGAGTGCATCCAGCGAAGGACAGAGCTTGGCGTG 169
QY 121 CTCGCTGGGCTGTCTGGAGTGCAGATGAGGGGACAATGATGACAGAGGCC 180
DB 170 CTCGCTGGGCTGTCTGGAGTGCAGATGAGGGGACAATGATGACAGAGGCC 229
QY 181 AGTTGAGCCACAGAGAGGCGCCGAGTGAAGAGAGCTCCACGGGGACACAGACAGAC 240
DB 230 AGTTGAGCCACAGAGAGGCGCCGAGTGAAGAGAGCTCCACGGGGACACAGACAGAC 289
QY 241 TTCGGGCAAGGATCCCAAGATTCACAGAGCAGAGAGAGAGGACAGACCTGACCTC 300
DB 290 TTCGGGCAAGGATCCCAAGATTCACAGAGCAGAGAGAGAGGACAGACCTGACCTC 349
QY 301 ATGTTACAGCTGTGAGGCGGAGAGATGACATCCGCTTGACAGCCAGCTGGAGGACACC 360
DB 350 ATGTTACAGCTGTGAGGCGGAGAGATGACATCCGCTTGACAGCCAGCTGGAGGACACC 409
QY 361 CGGCTTCCCGGCTCCGCTACCTGCTGTAGTTTCTACACGAGAGAGAGGCTGTAGC 420

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 729)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1009 row: c column: 18
High quality sequence stop: 686.

FEATURES

source

1. 729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4991561"
/cell_line="MG36"
/lab_host="DH10B"
/clone_lib="NIH_MGC_10"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
Location/Qualifiers
BASE COUNT 153 a 232 c 228 g 115 t 1 others
ORIGIN

Query Match 34.0%; Score 674; DB 12; Length 729;
Best Local Similarity 98.8%; Pred. No. 2.4e-121;
Matches 721; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
QY 1143 GGATGAGAGTGGCCCGCCAGCTGCTGCGCCACTGGAAGAGACGACCGCTTCATTGAGGC 1202
Db 1 GGATGAGAGTGGCCCGCCAGCTGCTGCGCCACTGGAAGAGACGACCGCTTCATTGAGGC 60
QY 1203 TGCAAGAGCAGAGGCGACCCACGCTGTGTCTCCACTGCAAGATGGCGCTACCGCGCTCAGC 1262
Db 61 TGCAAGAGCAGAGGCGACCCACGCTGTGTCTCCACTGCAAGATGGCGCTCAGCGCTCAGC 120
QY 1263 GGGCCACAGTGTGGCTATGCTGCAAGCAGTACGATGCAAGCCTGAGCAGCGCTGCG 1322
Db 121 GGGCCACAGTGTGGCTATGCTGCAAGCAGTACGATGCAAGCCTGAGCAGCGCTGCG 180
QY 1323 CCACGTGACAGAGCTCCGGCCCATGCGCCCGCCCAACCTGGCTTCTGCGCCAGCTGCA 1382
Db 181 CCACGTGACAGAGCTCCGG-CCATGCGCCCGCCCAACCTGGCTTCTGCGCCAGCTGCA 239
QY 1383 GATCTACCAAGGCGCATCTGACGGCGCAGCGCCAGAGCCATGTCTGGAGCAGAAAGTGGG 1442
Db 240 GATCTACCAAGGCGCATCTGACGGCGCAGCGCCAGAGCCATGTCTGGAGCAGAAAGTGGG 299
QY 1443 TGGGGTCTCCCGAGAGGAGCAGCCCGCTGAAGTCTCTACACCATTCCTCCCTCTTCC 1502
Db 300 TGGGGTCTCCCGAGAGGAGCAGCCCGCTGAAGTCTCTACACCATTCCTCCCTCTTCC 359
QY 1503 GCCAGAACCTGAGGGTGTGGGAGGAGAGAGTGTGAGGATGGA-AGAGAGCCAGGAG 1561
Db 360 GCCAGAACCTGAGGGTGTGGGAGGAGAGAGTGTGAGGATGGAAGAGAGCCAGGAG 419
QY 1562 CCCCCGAAAGAGAGCCTGGGCGACGGCCAGTATAACCTCCGAGGGTCTAGAGTCCA 1621
Db 420 CCCCCGAAAGAGAGCCTGGGCGACGGCCAGTATAACCTCCGAGGGTCTAGAGTCCA 479
QY 1622 TCAGTCTTCTGAGCCCTCC-TTGAAGCTGAGAGCAGCTCAGAGAGCAGTACATGCCA 1680
Db 480 TCAGTCTTCTGAGCCCTCCCTTTGAGAGCTGAGAGCAGCTCAGAGAGCAGTACATGCCA 539

QY 1681 GAGGTCTTCTTCCCGACGAGTCTTTCATGAGAGAGCCCTTGCAGCCCTTCCACAGCTT 1740
Db 540 GAGGTCTTCTTCCCGACGAGTCTTTCATGAGAGAGCCCTTGCAGCCCTTCCACAGCTT 599
QY 1741 GCAAGGACCAAGGAGGCGCCAGCAGGTGACAGAGGGGCGCTGAGCTTGAAGTCCGC 1800
Db 600 GCAAGGACCAAGGAGG-CAGCAGGTGACAGAGGGGCGCTGAGCTTGAAGTCCGC 658
QY 1801 CAGTCAGTGTATACCTCCAGGAGGAGTGCCTGTGTGCGCAACCGGACCGAGCTTCCAG 1860
Db 659 CAGTCAGTGTATACCTCCAGGAGGAGTGCCTGTGTGCGCAATCGAACCGAGCTTCCAG 718
QY 1861 GAGCAGGAGC 1870
Db 719 GAGCAGGAGC 728

RESULT 13

BE869222
LOCUS
DEFINITION 601445030F1 NIH_MGC_65 Homo sapiens CDNA clone IMAGE:3849342 5',
mRNA sequence.
ACCESSION BE869222 GI:10317998
VERSION BE869222
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 726)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM956 row: k column: 07
High quality sequence stop: 607.
Location/Qualifiers
source 1. 726
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3849342"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_65"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

FEATURES

source

BASE COUNT 137 a 235 c 226 g 128 t
ORIGIN
Query Match 34.0%; Score 673.8; DB 10; Length 726;
Best Local Similarity 98.9%; Pred. No. 2.7e-121;
Matches 689; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
QY 275 AGGAGCAGAGGAGCAGCAGCTGACCTCATGTGTACAGCTGTGAGGCGCGAGATGACATCC 334
Db 1 AGGAGCAGAGGAGCAGCAGCTGACCTCATGTGTGTACAGCTGTGAGGCGCGAGATGACATCC 60
QY 335 GCCTGGCAGCCAGCTGAGGAGCAGCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTTT 394
Db 61 GCCTGGCAGCCAGCTGAGGAGCAGCCCGGCTCCCGGCTCCGCTACCTGCTGTAGTTT 120
QY 395 CTACACGAGAGAGAGAGGTCTGAGCCAGGATGAGAGCGGTCTCTCTGGCGGTGATTTC 454

Db 121 CTACACGAGAAGAGGTCTGAGCCAGATGAGACGCTCTCTGGCGGTGATTTCC 180
QY 455 CTGACAGCAGCTCCCCAGCTGCACCTGGGCTGTCTTGGCCCTCTGAGTGACACCC 514
Db 181 CTGACAGCAGCTCCCCAGCTGCACCTGGGCTGTCTTGGCCCTCTGAGTGACACCC 240
QY 515 AGGTGTACTTAGATGAGACGCGGGCTTCAAGCTGACGTCTGGTGGGCAAGCCGATCT 574
Db 241 AGGTGTACTTAGATGAGACGCGGGCTTCAAGCTGACGTCTGGTGGGCAAGCCGATCT 300
QY 575 TCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGGTATTGACCAAGCATGTG 634
Db 301 TCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGGTATTGACCAAGCATGTG 360
QY 635 AGGCAGCTTAGGCACGCGGCTTGTACCGGAGTGACAGTGCCTTCACTGGGCCAGCCACT 694
Db 361 AGGCAGCTTAGGCACGCGGCTTGTACCGGAGTGACAGTGCCTTCACTGGGCCAGCCACT 420
QY 695 ACCAGAGAGACTGAATCTCCGAACAGAGCTGCTCAATGATGAGACGGCTATGGCCGACC 754
Db 421 ACCAGAGAGACTGAATCTCCGAACAGAGCTGCTCAATGATGAGACGGCTATGGCCGACC 480
QY 755 TGGAGTCTCTGCGGCTTCCAGCGCCGAGCTGGCGGCTCTCAGAACAGAGCAGATGG 814
Db 481 TGGAGTCTCTGCGGCTTCCAGCGCCGAGCTGGCGGCTCTCAGAACAGAGCAGATGG 540
QY 815 AGCAGCGATCCGTGTGAGCTGTGAAAGTGTGATGTCACTGAGAGAGTGTCA 874
Db 541 AGCAGCGATCCGTGTGAGCTGTGAAAGTGTGATGTCACTGAGAGAGTGTCA 600
QY 875 CTTCGAAGAGATCCGCGCAGGCTCTGAGCTGGCGCTGGCGCTCCCTCCAGCAGTACC 934
Db 601 CTTCGAAGAGATCCGCGCAGGCTCTGAGCTGGCGCTGGCGCTCCCTCCAGCAGTACC 660
QY 935 GTGACTTCATCGACAACAGATGCTGCTGTGTGGC 971
Db 661 GCGACTTCATCGACAACAGATGCTGCTGTGTGGC 696

RESULT 14
LOCUS BQ675874 952 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8058966 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6212874
5', mRNA sequence.

ACCESSION BQ675874
VERSION BQ675874.1 GI:21786708
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 952)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLCM2377 row: k column: 19
High quality sequence stop: 582.
Location/Qualifiers

FEATURES
source 1..952

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6212874"

/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_102"
/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH-MGC Library."

Query Match 33.9%; Score 670.8; DB 13; Length 952;
Best Local Similarity 94.7%; Pred. No. 1.1e-120;
Matches 739; Conservative 0; Mismatches 32; Indels 9; Gaps 4;

QY 1207 AGAGCAGAGGCGACCCAGCTGCTGTCTCACTGCAAGATGGCGCTCAGCGGCC 1266
Db 25 AGAGCAGAGGCGACCCAGCTGCTGTCTCACTGCAAGATGGCGCTCAGCGGCC 84
QY 1267 ACAGTGTGCTTATGCTATGAGCAGTACGATGACGCTTGAGAGAGCCCTGCGCAC 1326
Db 85 ACAGTGTGCTTATGCTATGAGCAGTACGATGACGCTTGAGAGAGCCCTGCGCAC 144
QY 1327 GTGACAGAGCTCCGGCCATCGCCCGCCCAACCTTGGCTTCTGCGCCAGCTGAGATC 1386
Db 145 GTGACAGAGCTCCGGCCATCGCCCGCCCAACCTTGGCTTCTGCGCCAGCTGAGATC 204
QY 1387 TACCAAGGCATCTGACCGCCAGCCCGCCAGAGCCATGTCTGGAGAGAAAGTGGTGG 1446
Db 205 TACCAAGGCATCTGACCGCCAGCCCGCCAGAGCCATGTCTGGAGAGAAAGTGGTGG 264
QY 1447 GTTCTCCCAAGAGAGAGCAGCCAGCCCTGAAAGTCTTAACCATTTCCACCTTTCCGCCA 1506
Db 265 GTTCTCCCAAGAGAGAGCAGCCAGCCCTGAAAGTCTTAACCATTTCCACCTTTCCGCCA 324
QY 1507 GAACCTGAGGGTGTGGGAGGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCCCG 1566
Db 325 GAACCTGAGGGTGTGGGAGGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCCCG 384
QY 1567 AAAGAAGAGCCTGGGCCAGGCCACGATATAAAGCTCCGAGGGTCAATGAGTCCATCAGT 1626
Db 385 AAAGAAGAGCCTGGGCCAGGCCACGATATAAAGCTCCGAGGGTCAATGAGTCCATCAGT 444
QY 1627 CTTCTGAGCCCTCTTGAAGCTGAGAGCAGCTCAGAGACCATGACATGCGAGAGTTC 1686
Db 445 CTTCTGAGCCCTCTTGAAGCTGAGAGCAGCTCAGAGACCATGACATGCGAGAGTTC 504
QY 1687 TTCTCTCCACAGAGTCTTACATGAAGAGCCTCTGACGCCCTTCCACAGCTTGCAAGG 1746
Db 505 TTCTCTCCACAGAGTCTTACATGAAGAGCCTCTGACGCCCTTCCACAGCTTGCAAGG 564
QY 1747 ACCAAGGAGGCGCAGAGTGAACAGGGGCTCAGCCTGCCCCGAGAGTCCCGCAGTCA 1806
Db 565 ACCAAGGAGGCGCAGAGTGAACAGGGGCTCAGCCTGCCCCGAGAGTCCCGCAGTCA 624
QY 1807 GTGTTACCTCTCAGGGCAGTGGC-GTGTGGCCAAACCGGACCGAGGCTTCCAGAGACA 1865
Db 625 GTGTTACCTCTCAGGGCAGTGGC-GTGTGGCCAAACCGGACCGAGGCTTCCAGAGAGC 684
QY 1866 GAGACAGGGGCGAGGGCA-GGGGCGAGAGAGCCCTGCA-TTCTCTTAAGCCAGGTTTC 1923
Db 685 GAGACAGGGGCGAGGGCA-GGGGCGAGAGAGCCCTGCA-TTCTCTTAAGCCAGGTTTC 744
QY 1924 C-----GGAAGTGTGTGACACAGGCCAGCTGATGACAGTGAAGAGAGGGCGAGGCC 1977
Db 745 CCGGAAAGGGGTGTGAGACAGGCCAGCCGTGATGACAAATTGGAAAAAGAGAGGGCC 804

RESULT 15
BG169297

LOCUS BG169297 910 bp mRNA linear EST 06-FEB-2001
DEFINITION 602321027F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4424070 5',
mRNA sequence.
ACCESSION BG169297
VERSION BG169297
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 910)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10167 row: f column: 07
High quality sequence stop: 742.

FEATURES
source location/Qualifiers
1..910
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4424070"
/tissue_type="hypernephroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_89"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 189 a 270 c 304 g 146 t 1 others
ORIGIN

Query Match 33.5%; Score 664; DB 10; Length 910;
Best Local Similarity 95.7%; Pred. No. 2.3e-119;
Matches 757; Conservative 0; Mismatches 26; Indels 8; Gaps 7;

QY 534 CGGGGGCTTCAGCGTGAAGTGGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCA 593
DB 1 CGGGGGCTTCAGCGTGAAGTGGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCA 60
QY 594 GACCATGTGGGCCACACTCCAGTATTGACCAAGCATGTGAGGAGCTCTAGGAGCGG 653
DB 61 GACCATGTGGGCCACACTCCAGTATTGACCAAGCATGTGAGGAGCTCTAGGAGCGG 119
QY 654 CCTTGTACCGGGTGGCAGTGCCTCACTGGGCCCACTATACAGAGAGACTGAATC 713
DB 120 CCTTGTACCGGGTGGCAGTGCCTCACTGGGCCCACTATACAGAGAGACTGAATC 179
QY 714 CGAACAGAGCTGCCTCAATGATGAGACGGCTATGCGGCACTGAGTCTTGCGGCTCC 773
DB 180 CGAACAGAGCTGCCTCAATGATGAGACGGCTATGCGGCACTGAGTCTTGCGGCTCC 239
QY 774 CAGCGCCGAGCTGGCGGCTCTCAGACAGAGAGATGAGACAGCGGATCCGTGCTGA 833
DB 240 CAGCGCCGAGCTGGCGGCTCTCAGACAGAGAGATGAGACAGCGGATCCGTGCTGA 299
QY 834 GCTGTGAAAGTGTGATGTCACTGAGAGTGTCACTTCAAAGAGATCCGCCA 893
DB 300 GCTGTGAAAGTGTGATGTCACTGAGAGTGTCACTTCAAAGAGATCCGCCA 358
QY 894 GGCTCTGAGCTGGCGCTGGCGCTCCCGCTCCAGAGTACCGGTGATTCATCGACAACCA 953
DB 359 GGCTCTGAGCTGGCGCTGGCGCTCCCGCTCCAGAGTACCGGTGATTCATCGACAACCA 418

QY 954 GATGCTGTGCTGTGGTGACAGCGGGACCGGAGCTCCCGCATCTTCCCCCACTCTACCT 1013
DB 419 GATGCTGTGCTGTGGTGACAGCGGGACCGGAGCTCCCGCATCTTCCCCCACTCTACCT 478
QY 1014 GGGCTCAGAGTGAACCGAGCAAACTTGAGAGAGCTGCAGAGAAAGAGGTCAACCAAT 1073
DB 479 GGGCTCAGAGTGAACCGAGCAAACTTGAGAGAGCTGCAGAGAAAGAGGTCAACCAAT 538
QY 1074 CTTGAACATGCCCCGGAGATGACAACTTCACTGAGCGCTTCACTCAACCAATGT 1133
DB 539 CTTGAACATGCCCCGGAGATGACAACTTCACTGAGCGCTTCACTCAACCAATGT 598
QY 1134 GCGCTCTGGGATGAGAGTGGGCCCGAGCTGTGCGCACTGGAAGAGAGCGACCGCTT 1193
DB 599 GCG-CTCTGGGATGAGAGTGG-CCAGCTGTGC--GACTGGAAGAGAGCGACCGCTT 654
QY 1194 CATTGA-GGCTGCAAGAGACAGGGCACCCACGCTGTGTCCACTGCAAGATGGCGTCA 1252
DB 655 CATGAGAGGCTGCAAGAGACAGGGCACCCACGCTGTGTCCACTGCAAGATGGCGTCA 714
QY 1253 GCCGCTCAGCGGCCACAGTGTGCTGCTATGCCATGAAGCAGTACGAATGAGCGCTGAGC 1312
DB 715 GCCGCTCAGCGG-CACAGGGCTGCTATGCCATGAAGCAGTACGAATGAGCGCTGAGC 773
QY 1313 AGGCCCTGGCC 1323
DB 774 GGCTTGGGCAC 784

Search completed: January 15, 2004, 18:29:18
Job time : 3788 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 13:11:13 : Search time 532 Seconds
(without alignments)
10046.781 Million cell updates/sec

Title: US-09-955-732-1
Perfect score: 1980
Sequence: 1 atggcctgtgtcacagttag.....gagaggagggcgcgcctga 1980

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1980	100.0	2718	24	AAD36061	Human dual-specific
2	1978.4	99.9	2781	22	AAF30485	Human protein phos
3	1978.4	99.9	2852	24	ABQ73250	Human MAP kinase p
4	1774	89.6	2618	24	AAD36063	Murine dual-specif
5	1774	89.6	2704	24	ABQ73249	Human MAP kinase p
6	1774	89.6	2704	24	ABQ73252	Human MAP kinase p
7	1684	85.1	2322	24	ABU40805	Human MAP kinase p
8	1345.6	68.0	2540	24	ABQ73251	Human MAP kinase p

9	1026	51.8	1026	22	AAF63578	Human phosphatase
10	851	43.0	2061	24	ABN59832	Novel human coding
11	694	35.1	1755	22	AAH14722	Human CDNA sequenc
12	694	35.1	1755	24	ABL40801	Human MAP kinase p
13	491	24.8	599	24	ABL40803	Human MAP kinase p
14	397	20.1	409	24	ABL40802	Human MAP kinase p
15	377.2	19.1	717	22	AAH98183	Human EST-derived
16	362	18.3	6374	22	AAD09491	Human SGP006 phosp
17	356.4	18.0	494	24	ABL40804	Human MAP kinase p
18	340.4	17.2	1949	22	AAD12965	Human dual-specific
19	329	16.6	426	24	ABL40800	Human MAP kinase p
20	308.8	15.6	1711	22	AAD12966	Human dual-specific
21	300.8	15.2	3488	24	ABL57466	Human protein phos
22	293.6	14.8	4467	23	ABL10739	Human Drosophila mel
23	282.8	14.3	2260	22	AAD09493	Human SGP001 phosp
24	279.4	14.1	1771	22	AAH22966	Human phosphatase
25	278	14.0	1348	22	AAH99712	Human protein enco
26	233.4	11.8	571	22	AAH07057	Human CDNA clone (
27	204	10.3	1052	22	AAD12967	Human dual-specific
28	163.4	8.3	969	22	AAS59843	Human novel cytol
29	163	8.2	5029	25	ABX34583	Human mdt CDNA SE
30	156.6	7.9	1450	22	AA541387	Human encoding nove
31	156.6	7.9	1450	22	AA534834	CDNA encoding nove
32	142	7.2	8002	23	ABL10738	Drosophila melanog
33	123.2	6.2	447	21	AACT75065	Human ORFX ORF620
34	101.8	5.1	1390	22	AA512700	Human CDNA encodin
35	101.8	5.1	1605	25	ABA00716	Human KPP-1 CDNA,
36	98.8	5.0	1236	24	AAD24021	Human protein phos
37	98.8	5.0	1357	24	ABQ55030	Human ovarian anti
38	98.8	5.0	1685	25	ABX62935	Human activated T
39	98.8	5.0	1709	24	ABK83823	Human CDNA differe
40	98.8	5.0	1710	25	ABX10959	CDNA sequence enco
41	93.8	4.7	687	22	AAF63575	Human phosphatase
42	89.2	4.5	2139	25	ABX09053	Human dual specif
43	89.2	4.5	2400	24	AA594769	Human DNA sequence
44	89.2	4.5	2479	25	ABX09060	Human dual specif
45	88	4.4	1435	25	ABV74143	Human dual specif

ALIGNMENTS

RESULT 1	
AAD36061	
ID	AAD36061 standard; CDNA; 2718 BP.
XX	
AC	AAD36061;
XX	
DT	09-AUG-2002 (first entry)
XX	
DE	Human dual-specificity phosphatase 15 (DSP-15) CDNA.
XX	
KW	Human; dual-specificity phosphatase 15; DSP-15; antiallergic; cyostatic;
KW	immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW	signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW	cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW	metabolic disease; allergy; screening; chromosome 11q; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	35..2014
FT	/*tag= a
XX	/product= "Human DSP-15 protein"
PN	WO200224740-A2.
XX	
PD	28-MAR-2002.
XX	
PF	19-SEP-2001; 2001WO-US29406.
XX	
PR	19-SEP-2000; 2000US-233833P.
PR	18-SEP-2001; 2001US-0955732.

XX (CEPT-) CEPTYR INC.
XX Luche RM, Wei B;
XX MPI; 2002-394127/42.
DR P-PSDB; AAE22729.
XX
PT New dual-specificity phosphatase 15 polypeptide and polynucleotides,
PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
PT diseases

PS Claim 7; Fig 1; 91pp; English.

XX The invention relates to a new isolated dual-specificity phosphatase 15
CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
CC residues. DSP-15 polypeptides may be used to identify agents that
CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
CC polypeptides, modulating agents, and/or polynucleotides encoding the
CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
CC abnormal cell growth, abnormal cell proliferation and cell cycle
CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
CC assays for modulators of enzyme activity and/or substrate binding. The
CC present sequence is human DSP-15 cDNA. Human DSP-15 gene is located on
CC chromosome 11q.

XX Sequence 2718 BP; 559 A; 893 C; 795 G; 471 T; 0 other;

Query Match 100.0%; Score 1980; DB 24; Length 2718;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTGGTCACTGAGCCGTTGCCCCCGGAGCGGCGCTCCACGCGCGTGGG 60
DB 35 ATGGCCCTGGTCACTGAGCCGTTGCCCCCGGAGCGGCGCTCCACGCGCGTGGG 94
QY 61 CCCTGGGAGCAGGCGGTCAGCGAAGAGTCACTCCAGCGAAGCAGAGCTTTGGCGTG 120
DB 95 CCCTGGGAGCAGGCGGTCAGCGAAGAGTCACTCCAGCGAAGCAGAGCTTTGGCGTG 154
QY 121 CTCCGTGGGCTGTCTCTGGGACTGAGATGAGGGGCAATGATGATGACAGAGGCC 180
DB 155 CTCCGTGGGCTGTCTCTGGGACTGAGATGAGGGGCAATGATGATGACAGAGGCC 214
QY 181 AGTTCTGAGCAACAGAGAGGCGCCGAGTGAAGAGAGCTCCAGCGGACAGACAGAC 240
DB 215 AGTTCTGAGCAACAGAGAGGCGCCGAGTGAAGAGAGCTCCAGCGGACAGACAGAC 274
QY 241 TTGGGCAAGATCCAGAGTCCCGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 275 TTGGGCAAGATCCAGAGTCCCGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAG 334
QY 301 ATGTACAGCTGTGAGCGCGAGATGACATCCGCTGGCAGCGCCAGCTGAGCAACC 360
DB 335 ATGTACAGCTGTGAGCGCGAGATGACATCCGCTGGCAGCGCCAGCTGAGCAACC 394
QY 361 CGGCTCCCGGCTCCGCTACCTGCTGTAGTTTCTACAGAGAGAGAGAGAGAGAG 420
DB 395 CGGCTCCCGGCTCCGCTACCTGCTGTAGTTTCTACAGAGAGAGAGAGAGAGAG 454
QY 421 CAGATGAGACGGTCTCTCTGGGCGTGAATTTCCCTGACAGAGAGAGAGAGAGAG 480
DB 455 CAGATGAGACGGTCTCTCTGGGCGTGAATTTCCCTGACAGAGAGAGAGAGAGAG 514
QY 481 CTGGGCTGTCTTGGGCTCTGAGATGACACCCAGAGAGAGAGAGAGAGAGAGAG 540
DB 515 CTGGGCTGTCTTGGGCTCTGAGATGACACCCAGAGAGAGAGAGAGAGAGAGAG 574

QY 541 TTCAGCGTACGCTGTGGTGGCAAGCGGATCTCTCAAGCCATCTCCATCCAGACCATG 600
DB 575 TTCAGCGTACGCTGTGGTGGCAAGCGGATCTCTCAAGCCATCTCCATCCAGACCATG 634
QY 601 TGGCCACACTCCAGGATTTGACCAAGCATGTGAGGAGCTCTAGGAGCGGCTGTGA 660
DB 635 TGGCCACACTCCAGGATTTGACCAAGCATGTGAGGAGCTCTAGGAGCGGCTGTGA 694
QY 661 CCGGCTGAGTGCCTTCACTGGGCGAGCCACTACAGAGAGAGAGAGAGAGAGAGAG 720
DB 695 CCGGCTGAGTGCCTTCACTGGGCGAGCCACTACAGAGAGAGAGAGAGAGAGAGAG 754
QY 721 AGCTGCTCAATGATGAGACGGCTATGCGCCGACCTGAGTCTCTGCGCTCCAGCGCC 780
DB 755 AGCTGCTCAATGATGAGACGGCTATGCGCCGACCTGAGTCTCTGCGCTCCAGCGCC 814
QY 781 GAGCTGGCGGCTCTCAGAACAGAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 815 GAGCTGGCGGCTCTCAGAACAGAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAG 874
QY 841 AAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 875 AAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 934
QY 901 GAGCTGGCGGCTCTCAGAACAGAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 935 GAGCTGGCGGCTCTCAGAACAGAGCAGATGAGAGAGAGAGAGAGAGAGAGAGAG 994
QY 961 CTGCTGTGGGCAAGCGGAGCCGAGCTCCGATCTTCCCGACCTTACCTTGGGCTCA 1020
DB 995 CTGCTGTGGGCAAGCGGAGCCGAGCTCCGATCTTCCCGACCTTACCTTGGGCTCA 1054
QY 1021 GAGTGAACGAGCAAAACCTGGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1055 GAGTGAACGAGCAAAACCTGGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAG 1114
QY 1081 ATGGCCCGGAGATTGACAACTTCTAACCCTGAGCGCTTCACTACCAATATGCGCTC 1140
DB 1115 ATGGCCCGGAGATTGACAACTTCTAACCCTGAGCGCTTCACTACCAATATGCGCTC 1174
QY 1141 TGGATGAGAGAGTGGGCCAGCTGTGCGCACTGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1175 TGGATGAGAGAGTGGGCCAGCTGTGCGCACTGAGAGAGAGAGAGAGAGAGAGAG 1234
QY 1201 GCTGCAAGAGACAGGGCACCCAGCTGTGCTCACTGCAAGATGGCGTACCGCTCA 1260
DB 1235 GCTGCAAGAGACAGGGCACCCAGCTGTGCTCACTGCAAGATGGCGTACCGCTCA 1294
QY 1261 GCGGCCACAGTCTGCGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1295 GCGGCCACAGTCTGCGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1354
QY 1321 CGCCACGTGAGAGAGTCCCGGCCATCGCCCGCCCAACCTTGCTTCTGCGCCAGCTG 1380
DB 1355 CGCCACGTGAGAGAGTCCCGGCCATCGCCCGCCCAACCTTGCTTCTGCGCCAGCTG 1414
QY 1381 CAGATCTACAGAGGATCTCTGACGCGCCAGCGCCAGAGCCATGTCTGGAGAGAGAA 1440
DB 1415 CAGATCTACAGAGGATCTCTGACGCGCCAGCGCCAGAGCCATGTCTGGAGAGAGAA 1474
QY 1441 GGTGGGCTTCCCGAGAGAGACCCAGCCCTGAACTCTTACACCATTTCCACCTCTT 1500
DB 1475 GGTGGGCTTCCCGAGAGAGACCCAGCCCTGAACTCTTACACCATTTCCACCTCTT 1534
QY 1501 CCGCCAGAACTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 1535 CCGCCAGAACTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1594
QY 1561 GCCCGGAAGAGAGCGCTGGGCCAGCGCCAGATATAACCTCCAGGGGCTATGAGTCC 1620
DB 1595 GCCCGGAAGAGAGCGCTGGGCCAGCGCCAGATATAACCTCCAGGGGCTATGAGTCC 1654

QY	1621	ATCAGTCTTCTTGAGAGCCCTCTCTTGAGCTGGAGAGCACTCAGAGACCAGTGCATGCGCA	1680
Db	1655	ATCAGTCTTCTTGAGAGCCCTCTCTTGAGCTGGAGAGCACTCAGAGACCAGTGCATGCGCA	1714
QY	1681	GAGGCTTCTCTTCCCAACAGTCTTCACATGAAGAGCCTCTGCAGCCCTTCCCAAGCTT	1740
Db	1715	GAGGCTTCTCTTCCCAACAGTCTTCACATGAAGAGCCTCTGCAGCCCTTCCCAAGCTT	1774
QY	1741	GCAAGGACCAAGGGAGGCCAGAGTGGACAGAGGGGCTCAGCCTGCCCTGAAGTCCGC	1800
Db	1775	GCAAGGACCAAGGGAGGCCAGAGTGGACAGAGGGGCTCAGCCTGCCCTGAAGTCCGC	1834
QY	1801	CAGTCAGTGGTTACCTCCAGGGCAGTGCCGTGCTGGCCAAACCGAACCCAGGCGCTTCCAG	1860
Db	1835	CAGTCAGTGGTTACCTCCAGGGCAGTGCCGTGCTGGCCAAACCGAACCCAGGCGCTTCCAG	1894
QY	1861	GAGCAGGAGCAGGGGCGAGGGGCGAGGGGCGAGAGGCCCTGCATTTCCTCTACGCCCAGG	1920
Db	1895	GAGCAGGAGCAGGGGCGAGGGGCGAGGGGCGAGAGGCCCTGCATTTCCTCTACGCCCAGG	1954
QY	1921	TTCCGGAAGGTGCTGAGACAGGCCCAGCGTGCATGACAGTGGAGAGGAGGGCGAGGCGCTGA	1980
Db	1955	TTCCGGAAGGTGCTGAGACAGGCCCAGCGTGCATGACAGTGGAGAGGAGGGCGAGGCGCTGA	2014

RESULT 2	
AAAF30485	
ID	AAAF30485 standard; cDNA; 2781 BP.
AC	AAAF30485;
DT	29-MAY-2001 (first entry)
DE	Human protein phosphatase and kinase protein-10 cDNA 5039718CB1.
KW	Protein phosphatase and kinase protein; PPKP-10; human;
KW	gastrointestinal disorder; immune system disorder;
KW	neurological disorder; cell proliferative disorder; cancer;
KW	diagnosis; therapy; ss.
OS	Homo sapiens.
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	87..2066
FT	/tag= a
FT	87..143
FT	/tag= b
FT	144..2063
FT	/tag= c
FT	219..279
FT	/tag= d
FT	/note= "unique fragment"
FT	921..980
FT	/tag= e
FT	/note= "unique fragment"
PN	WO200120004-A2.
PD	22-MAR-2001.
PF	14-SEP-2000; 2000WO-US25515.
PR	15-SEP-1999; 99US-0154141.
PA	(INCY-) INCYTE GENOMICS INC.
PI	Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y,
PI	Lu DAM;
DR	WPI; 2001-244811/25.
DR	P-PADB; AAB20331.
XX	Novel human protein phosphatase and kinase proteins for diagnosis

PT	treatment and prevention of gastrointestinal, immune system,
PT	neurological and cell proliferative disorders -
XX	
XX	Claim'5; Page 101-102; 103pp; English.
XX	
CC	The present sequence is that of cDNA encoding novel human
CC	protein phosphatase and kinase protein PPHK-10 (see AAB20331).
CC	The cDNA was initially identified in Incyte Clone ID No. 5039718CB1,
CC	from a colon tumour tissue library. Tissues that express PPHK-10
CC	(as a fraction of total tissues expressing PPHK-10) include
CC	reproductive (0.343), gastrointestinal (0.194) and haematopoietic
CC	or immune (0.134). Diseases or conditions associated with tissues
CC	expressing PPHK-10 (as a fraction of total tissues expressing
CC	PPHK-10) include cancer (0.552), inflammation or trauma (0.314) or
CC	cell proliferation (0.090). The encoded protein shows homology to
CC	Drosophila melanogaster MAP kinase phosphatase. The invention
CC	provides human PPHK-1 to -11 polypeptides (see AAB20322-32) and
CC	polynucleotides (see AAF30476-86). It also provides expression
CC	vectors, host cells, antibodies, agonists and antagonists, as well
CC	as methods for diagnosing, treating or preventing disorders
CC	associated with expression of PPHK, including gastrointestinal
CC	disorders, immune system disorders, neurological disorders and cell
CC	proliferative disorders, including cancer.
XX	
XX	
Sequence	2781 BP; 576 A; 906 C; 820 G; 479 T; 0 other;
Query Match	99.9%; Score 1978.4; DB 22; Length 2781;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1979; Conservative	0; Mismatches 1; Indels 0; Gaps 0
QY	1 ATGCCCTTGTCACAGTAGAGCCGTTGCCCGGGGAGCGGCGCTCCACGCCCTGGGG 60
DB	87 ATGCCCTTGTCACAGTAGAGCCGTTGCCCGGGGAGCGGCGCTCCACGCCCTGGGG 146
QY	61 CCCTGGACCAAGCGGTCCAGGAGAGAGTGCACCTCAGCGAAGCAGAGCTTGGGTG 120
DB	147 CCCTGGACCAAGCGGTCCAGGAGAGAGTGCACCTCAGCGAAGCAGAGCTTGGGTG 206
QY	121 CTCCTGGGCTGTCTCTGGGACTCAGAGATGAGGGGCAATGATGACAGAGGCC 180
DB	207 CTCCTGGGCTGTCTCTGGGACTCAGAGATGAGGGGCAATGATGACAGAGGCC 266
QY	181 AGTTCTGAGCCAACAGAGAGGCCCGGAGTGAGGAGAGCTCCACGGGGACAGACAG 240
DB	267 AGTTCTGAGCCAACAGAGAGGCCCGGAGTGAGGAGAGCTCCACGGGGACAGACAG 326
QY	241 TTCGGGCAAGATCCCAAGTCCCAAGAGCAGAGAGAGAGGAGCAGCCTGCACCTC 300
DB	327 TTCGGGCAAGATCCCAAGTCCCAAGAGCAGAGAGAGAGGAGCAGCCTGCACCTC 386
QY	301 ATGTACAGCTGTGAGCGCGAGAGATGACATCCGCTGGAGCCCACTGGAGGACCC 360
DB	387 ATGTACAGCTGTGAGCGCGAGAGATGACATCCGCTGGAGCCCACTGGAGGACCC 446
QY	361 CGGCTCCCGGGCTCCGCTACTGTGTAGTTTCTACACGAGAAGAGAAGTCTGAGC 420
DB	447 CGGCTCCCGGGCTCCGCTACTGTGTAGTTTCTACACGAGAAGAGAAGTCTGAGC 506
QY	421 CAGATGAGACGGTCTCTCTGGGGCTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACC 480
DB	507 CAGATGAGACGGTCTCTCTGGGGCTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACC 566
QY	481 CTGGGCTGTCTTGCCCTCTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGGC 540
DB	567 CTGGGCTGTCTTGCCCTCTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGGC 626
QY	541 TTCAGCGTGACGTCTGGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
DB	627 TTCAGCGTGACGTCTGGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 686
QY	601 TGGGCCACACTCCAGGTATTGCAACCAAGCATGTGAGGAGAGCTTAGGAGCGGCTTGTA 660
DB	687 TGGGCCACACTCCAGGTATTGCAACCAAGCATGTGAGGAGAGCTTAGGAGCGGCTTGTA 746

```
QY 661 CCGGTGGAGTGCCTCACTGGCCAGCCACTACAGAGAGACTGAATCCGAACAG 720
    |||
Db 747 CCGGTGGAGTGCCTCACTGGCCAGCCACTACAGAGAGACTGAATCCGAACAG 806
QY 721 AGCTGCTCAATGAGTGGACGGCTATGGCCGACCTGGAGTCTCTGGCCCTCCAGCCGC 780
    |||
Db 807 AGCTGCTCAATGAGTGGACGGCTATGGCCGACCTGGAGTCTCTGGCCCTCCAGTGC 866
QY 781 GAGCTGGCGGGTCTCTAGAAACAGAGCAGATGGAGAGCGGATCCGTGCTGAGCTGTGG 840
    |||
Db 867 GAGCTGGCGGGTCTCTAGAAACAGAGCAGATGGAGAGCGGATCCGTGCTGAGCTGTGG 926
QY 841 AAAGTGTGATGTCACTGACCTGGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 900
    |||
Db 927 AAAGTGTGATGTCACTGACCTGGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 986
QY 901 GAGCTGGCGCTGGGGCTCCCTCTCAGCAGTACCCTGAATTCTCATCGACCAACAGATGCTG 960
    |||
Db 987 GAGCTGGCGCTGGGGCTCCCTCTCAGCAGTACCCTGAATTCTCATCGACCAACAGATGCTG 1046
QY 961 CTGCTGGTGGCAGACGGGGACCGACCTCCCGCATCTTCCCGCCTCTAAGCTGGCTCA 1020
    |||
Db 1047 CTGCTGGTGGCAGACGGGGACCGACCTCCCGCATCTTCCCGCCTCTAAGCTGGCTCA 1106
QY 1021 GAGTGAACGACCAAACTGGAGAGCTGCAGAGAGAACAGGGTCAACCCATCTTGAAC 1080
    |||
Db 1107 GAGTGAACGACCAAACTGGAGAGCTGCAGAGAGAACAGGGTCAACCCATCTTGAAC 1166
QY 1081 ATGCCCCGGGAGATTGACAACTTCTACCTGAGCGCTTCACTTACCACTATGTGGCCTC 1140
    |||
Db 1167 ATGCCCCGGGAGATTGACAACTTCTACCTGAGCGCTTCACTTACCACTATGTGGCCTC 1226
QY 1141 TGGATGAGAGTGGCCAGCTGTGCGCAGCTGCGCAGTGAAGAGAGACGACCGCTTATTGAG 1200
    |||
Db 1227 TGGATGAGAGTGGCCAGCTGTGCGCAGCTGCGCAGTGAAGAGAGACGACCGCTTATTGAG 1286
QY 1201 GCTGCAAGAGCAGGGCACCAGCTGTGCTGCTCACTGCAAGATGGCGCTGACCGCTCA 1260
    |||
Db 1287 GCTGCAAGAGCAGGGCACCAGCTGTGCTGCTCACTGCAAGATGGCGCTGACCGCTCA 1346
QY 1261 GCGGCCACAGTGTGGCTATGCCATGAAGCAGTACGATGCAAGCTGGAGCAGGCCCTG 1320
    |||
Db 1347 GCGGCCACAGTGTGGCTATGCCATGAAGCAGTACGATGCAAGCTGGAGCAGGCCCTG 1406
QY 1321 CGCCACGTGACGAGACTCCGGCCCATCGCCGCCCAACCTGGCTTCTCGGCCAGCTG 1380
    |||
Db 1407 CGCCACGTGACGAGACTCCGGCCCATCGCCGCCCAACCTGGCTTCTCGGCCAGCTG 1466
QY 1381 CAGATCTACAGGGCATCTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCAGAAAGTG 1440
    |||
Db 1467 CAGATCTACAGGGCATCTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCAGAAAGTG 1526
QY 1441 GGTGGGTCTCCCAAGAGAGCAGCCAGCCCTGAAAGTCTTACACCATTTCCACCTCTT 1500
    |||
Db 1527 GGTGGGTCTCCCAAGAGAGCAGCCAGCCCTGAAAGTCTTACACCATTTCCACCTCTT 1586
QY 1501 CCGCCAGAACCTGAGGGTGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGGCA 1560
    |||
Db 1587 CCGCCAGAACCTGAGGGTGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGGCA 1646
QY 1561 GCGCCGAAAGAGAGCCTGGGGCCAGGGCCACGTATAACCTCCGAGGGGTATGAGGTCC 1620
    |||
Db 1647 GCGCCGAAAGAGAGCCTGGGGCCAGGGCCACGTATAACCTCCGAGGGGTATGAGGTCC 1706
QY 1621 ATCAGTCTTCTGGAGCCCTCTTGGAGCTGAGAGACACTCAGAGACCATGATGCCA 1680
    |||
Db 1707 ATCAGTCTTCTGGAGCCCTCTTGGAGCTGAGAGACACTCAGAGACCATGATGCCA 1766
QY 1681 GAGGTCTTCTTCCACGAGTCTTCAATGAAGAGCCTCTGACGCCCTTCCACAGCTT 1740
    |||
Db 1767 GAGGTCTTCTTCCACGAGTCTTCAATGAAGAGCCTCTGACGCCCTTCCACAGCTT 1826
```

```
QY 1741 GCAAGACCAAGGAGGCCAGAGGTGACAGAGGGGCTCAGCCTGCTGAAGTCCCGC 1800
    |||
Db 1827 GCAAGACCAAGGAGGCCAGAGGTGACAGAGGGGCTCAGCCTGCTGAAGTCCCGC 1886
QY 1801 CAGTCACTGTTACCTCCAGGGCAGTCCCGTGTGTGCCCAACCGACCGCCTTCCAG 1860
    |||
Db 1887 CAGTCACTGTTACCTCCAGGGCAGTCCCGTGTGTGCCCAACCGACCGCCTTCCAG 1946
QY 1861 GAGCAGACAGAGGGGAGAGGGGAGAGAGCCCTGCATTTCTCTACGCCCAGG 1920
    |||
Db 1947 GAGCAGACAGAGGGGAGAGGGGAGAGAGCCCTGCATTTCTCTACGCCCAGG 2006
QY 1921 TTCCGGAAGTGTGAGACAGGCCACCGCTGCATGACGTGAGAGAGAGGGCGAGCCTGA 1980
    |||
Db 2007 TTCCGGAAGTGTGAGACAGGCCACCGCTGCATGACGTGAGAGAGAGGGCGAGCCTGA 2066
```

```
RESULT 3
AB073250
ID AB073250 standard; cDNA; 2852 BP.
XX
AC AB073250;
XX
DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 2 cDNA sequence SEQ ID NO:2.
XX
KW Human; phosphatase; mitogen activated protein kinase phosphatase;
XX MAP kinase; enzyme; chromosome 11; gene; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..56
FT CDS /*tag= a
FT /*tag= b
FT /*product= "MAP kinase phosphatase splice form 2"
FT 3'UTR 2037..2852
FT /*tag= c
XX
PN WO200242436-A2.
XX
PD 30-MAY-2002.
XX
PE 07-NOV-2001; 2001WO-US42995.
XX
PR 20-NOV-2000; 2000US-0715177.
PR 18-JAN-2001; 2001US-0761640.
XX
PA (PEKE ) PE CORP NY.
XX
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-575237/61.
DR P-PSDB; ABP51654.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies -
XX
PS Claim 1; Fig 1B; 85bp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
```

CC encodes human MAP kinase phosphatase splice form 2 from the present
XX invention.

Sequence 2852 BP; 674 A; 895 C; 807 G; 476 T; 0 other;

Query Match 99.9%; Score 1978.4; DB 24; Length 2852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGCCCTGTACAGTAGAGCCGTTGCCCCCGGCGACGGCGCTCCAGCCCGTGGG 60
    |||
Db 57 ATGGCCCTGTACAGTAGAGCCGTTGCCCCCGGCGACGGCGCTCCAGCCCGTGGG 116
QY 61 CCTGGGACCGAGCGGTCCAGCGAAGAGTGAAGTCCAGCGAAGGAGAGCTTTGGGTG 120
    |||
Db 117 CCTGGGACCGAGCGGTCCAGCGAAGAGTGAAGTCCAGCGAAGGAGAGCTTTGGGTG 176
QY 121 CTCCTGGGGCTGTCTGGGACTGACAGATGAGGGGACAATGATGATGACAGAGGCC 180
    |||
Db 177 CTCCTGGGGCTGTCTGGGACTGACAGATGAGGGGACAATGATGATGACAGAGGCC 236
QY 181 AGTTCTGAGCCCAACAGAGAGGCCCGAGTGAAGAGAGCTCCACGGGGACCAGACAG 240
    |||
Db 237 AGTTCTGAGCCCAACAGAGAGGCCCGAGTGAAGAGAGCTCCACGGGGACCAGACAG 296
QY 241 TTCGGGCAAGGATCCCAAGAGTCCCAAGAAAGCAAGAGAGAGAGGACCACTGACCTC 300
    |||
Db 297 TTCGGGCAAGGATCCCAAGAGTCCCAAGAAAGCAAGAGAGAGGACCACTGACCTC 356
QY 301 ATGCTACAGCTGTCTGAGGCGGCAAGATGACATCCGCTTGGCAGCCCACTGAGGACA 360
    |||
Db 357 ATGCTACAGCTGTCTGAGGCGGCAAGATGACATCCGCTTGGCAGCCCACTGAGGACA 416
QY 361 CGGCTCTCCCGGCTCCGCTACTGCTGTAGTTCTACACGAGAGAGAGAGGTCTGAGC 420
    |||
Db 417 CGGCTCTCCCGGCTCCGCTACTGCTGTAGTTCTACACGAGAGAGAGAGGTCTGAGC 476
QY 421 CAGGATGAGACGGTCTCTGAGGCGGCTGATTTCCCTGACAGCAGCTCCCCAGCTGCAC 480
    |||
Db 477 CAGGATGAGACGGTCTCTGAGGCGGCTGATTTCCCTGACAGCAGCTCCCCAGCTGCAC 536
QY 481 CTGGCCCTGTCTTGTCCCTCTGAGTGAACACACCGGTGTACTTATGATGAGACGGGG 540
    |||
Db 537 CTGGCCCTGTCTTGTCCCTCTGAGTGAACACACCGGTGTACTTATGATGAGACGGGG 596
QY 541 TTCAGCGTGAAGTGTGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 600
    |||
Db 597 TTCAGCGTGAAGTGTGTGGGCAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 656
QY 601 TGGGCAACACTCCAGGTATGCAACCAAGCATGTGAGGAGAGCTTAGGAGCGGCTTGT 660
    |||
Db 657 TGGGCAACACTCCAGGTATGCAACCAAGCATGTGAGGAGAGCTTAGGAGCGGCTTGT 716
QY 661 CCGGTTGGCAGTGTCCCTCACTGGGGCCAGCCACTACAGAGAGAGACTGAATCCGAACAG 720
    |||
Db 717 CCGGTTGGCAGTGTCCCTCACTGGGGCCAGCCACTACAGAGAGAGACTGAATCCGAACAG 776
QY 721 AGCTGCTCATGATGATGAGCGGCTATGCGGCACTGGAAGTCTTGCGGCTCCAGCGCC 780
    |||
Db 777 AGCTGCTCATGATGATGAGCGGCTATGCGGCACTGGAAGTCTTGCGGCTCCAGCGCC 836
QY 781 GAGCTGGCGGGTCTCTCAAGACAGAGCAGATGAGCAGGCGATCCGTGCTGAGCTGTGG 840
    |||
Db 837 GAGCTGGCGGGTCTCTCAAGACAGAGCAGATGAGCAGGCGATCCGTGCTGAGCTGTGG 896
QY 841 AAAGTGTGATGTCAAGTGAAGTGAAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 900
    |||
Db 897 AAAGTGTGATGTCAAGTGAAGTGAAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 956
QY 901 GAGCTGGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAAACAGATGCTG 960
    |||
Db 957 GAGCTGGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAAACAGATGCTG 1016
XX
```

```
QY 961 CTGCTGGTGACACAGCGGGACCGAGCTCCCGCATCTTCCCCACCTTAACCTGGGCTCA 1020
    |||
Db 1017 CTGCTGGTGACACAGCGGGACCGAGCTCCCGCATCTTCCCCACCTTAACCTGGGCTCA 1076
QY 1021 GAGTGAACCGACGAACCTTGAGAGAGCTGACAGAGAAACAGGCTACCCACATCTTGAAC 1080
    |||
Db 1077 GAGTGAACCGACGAACCTTGAGAGAGCTGACAGAGAAACAGGCTACCCACATCTTGAAC 1136
QY 1081 ATGGCCCGGAGATTGACAACTTCTAACCTGAGCGCTTCACTACCAATGCGGCTC 1140
    |||
Db 1137 ATGGCCCGGAGATTGACAACTTCTAACCTGAGCGCTTCACTACCAATGCGGCTC 1196
QY 1141 TGGGATGAGAGTCCGCCAGCTGCTGCGGCACTGGAAGAGACGCAACGGCTTCATTGAG 1200
    |||
Db 1197 TGGGATGAGAGTCCGCCAGCTGCTGCGGCACTGGAAGAGACGCAACGGCTTCATTGAG 1256
QY 1201 GCTGCAAGACCAAGGGCAACCAAGTGTGTGCTGCTGCAAGATGGGCGTACCGGCTCA 1260
    |||
Db 1257 GCTGCAAGACCAAGGGCAACCAAGTGTGTGCTGCTGCAAGATGGGCGTACCGGCTCA 1316
QY 1261 GCGGCCACAGTGTGCTGCTTATGCCATGAAGCAGTACGAATGCAAGCTTGAGCAGGCC 1320
    |||
Db 1317 GCGGCCACAGTGTGCTGCTTATGCCATGAAGCAGTACGAATGCAAGCTTGAGCAGGCC 1376
QY 1321 CGCCACGTGACAGAGCTCCGGCCCATGCGCCGCCCAACCTGGCTTCTGGGCCAGCTG 1380
    |||
Db 1377 CGCCACGTGACAGAGCTCCGGCCCATGCGCCGCCCAACCTGGCTTCTGGGCCAGCTG 1436
QY 1381 CAGATCTACAGGGCATCTGACGGCCAGCGCCAGAGCCATGTCTGGAGAGAGAAAGTG 1440
    |||
Db 1437 CAGATCTACAGGGCATCTGACGGCCAGCGCCAGAGCCATGTCTGGAGAGAGAAAGTG 1496
QY 1441 GGTGGGTCTCCCAAGAGAGCAACCAAGCCCTGAAGTCTCTACACCATTCACACTCTT 1500
    |||
Db 1497 GGTGGGTCTCCCAAGAGAGCAACCAAGCCCTGAAGTCTCTACACCATTCACACTCTT 1556
QY 1501 CCGCAGAACTGAGGTTGGTGGGAGAGAGAGGTTGAGCATGGAAGAGAGCCAGGCA 1560
    |||
Db 1557 CCGCAGAACTGAGGTTGGTGGGAGAGAGAGGTTGAGCATGGAAGAGAGCCAGGCA 1616
QY 1561 GCCCCGAAGAAGAGAGCTGGGCCACGGCCACGATATAACCTCCGAGGGGTGATGAGTCC 1620
    |||
Db 1617 GCCCCGAAGAAGAGAGCTGGGCCACGGCCACGATATAACCTCCGAGGGGTGATGAGTCC 1676
QY 1621 ATCAGTCTTCTGAGAGCCCTCCTTGGAGCTGAGAGACACTCAGAGACCAGTACATGCCA 1680
    |||
Db 1677 ATCAGTCTTCTGAGAGCCCTCCTTGGAGCTGAGAGACACTCAGAGACCAGTACATGCCA 1736
QY 1681 GAGGTCTTCTTCCCAAGAGTCTTACATGAAGAGCCTCTGACGCCCTTCCACAGCTT 1740
    |||
Db 1737 GAGGTCTTCTTCCCAAGAGTCTTACATGAAGAGCCTCTGACGCCCTTCCACAGCTT 1796
QY 1741 GCAAGGACCAAGGAGGAGCCAGCAGGTGACAGAGGGGCTCAAGCTGCTGTAAGTCCCGC 1800
    |||
Db 1797 GCAAGGACCAAGGAGGAGCCAGCAGGTGACAGAGGGGCTCAAGCTGCTGTAAGTCCCGC 1856
QY 1801 CAGTCAAGTGTAACTCTCCAGGAGAGTCCGCTGTGAGCAACCGGACCCAGGCTTCCAG 1860
    |||
Db 1857 CAGTCAAGTGTAACTCTCCAGGAGAGTCCGCTGTGAGCAACCGGACCCAGGCTTCCAG 1916
QY 1861 GAGCAGAGAGAGGGGAGAGGGGAGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
    |||
Db 1917 GAGCAGAGAGAGGGGAGAGGGGAGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1976
QY 1921 TTCCGAAAGTGTGAGACAGGCAAGCGTGCATGACATGAGAGAGAGAGAGAGAGAGAG 1980
    |||
Db 1977 TTCCGAAAGTGTGAGACAGGCAAGCGTGCATGACATGAGAGAGAGAGAGAGAGAGAGAG 2036
XX
```

RESULT 4
AAD36063
ID AAD36063 standard; cDNA; 2618 BP.
XX

AC AAD36063;
XX 09-AUG-2002 (first entry)
DT
XX Murine dual-specificity phosphatase 15 (DSP-15) cDNA.
DE
XX Murine; dual-specificity phosphatase 15; DSP15; antiallergic; cytostatic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW metabolic disease; allergy; screening; gene; ss.
OS
XX Mus musculus.
FH Key Location/Qualifiers
FT CDS 35..1450
FT /*tag= a
FT /product= "Murine DSP-15 protein"
XX WO200224740-A2.
XX PD 28-MAR-2002.
XX 19-SEP-2001; 2001WO-US29406.
XX 19-SEP-2000; 2000US-233833P.
PR 18-SEP-2001; 2001US-0955732.
XX (CEPT-) CEPTYR INC.
XX Lucbe RM, Wei B;
XX WPI; 2002-394127/42.
DR P-PSDB; AAE22733.
XX New dual-specificity phosphatase 15 polypeptide and polynucleotides,
PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
PT diseases
XX Claim 56; Fig 4; 91pp; English.
PS
XX The invention relates to a new isolated dual-specificity phosphatase 15
CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
CC residues. DSP-15 polypeptides may be used to identify agents that
CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
CC polypeptides, modulating agents, and/or polynucleotides encoding the
CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
CC abnormal cell growth, abnormal cell proliferation and cell cycle
CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
CC assays for modulators of enzyme activity and/or substrate binding. The
CC present sequence is murine DSP-15 cDNA.
XX
SQ Sequence 2618 BP; 538 A; 857 C; 769 G; 454 T; 0 other;
Query Match 89.6%; Score 1774; DB 24; Length 2618;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1882; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

DB 155 CTCCTGGGGCTGTCTCTGGAGCTGACGATGAGAGGGAACATGATGATGACAGAGGCC 214
QY 181 AGTTCTGAGCCAAACAGAGAAGGCCCGAGTGAGAGAGCTCCACGGGACAGACAGAC 240
DB 215 AGTTCTGAGCCAAACAGAGAAGGCCCGAGTGAGAGAGCTCCACGGGACAGACAGAC 274
QY 241 TTCGGGCAAGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 275 TTCGGGCAAGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 334
QY 301 ATGCTACAGCTGTGAGAGGCCGAGATGACATCCGCTGAGAGAGAGAGAGAGAGAGAGAG 360
DB 335 ATGCTACAGCTGTGAGAGGCCGAGATGACATCCGCTGAGAGAGAGAGAGAGAGAGAG 394
QY 361 CGGCTCCCGGCTCCGCTACCTGCTGAGTATTCTTACAGAGAGAGAGAGAGAGAGAGAG 420
DB 395 CGGCTCCCGGCTCCGCTACCTGCTGAGTATTCTTACAGAGAGAGAGAGAGAGAGAGAG 454
QY 421 CAGGATGAGAGAGTCTCTCTGAGGCGGTGATTTCCCTGACAGAGAGTCCCGGAGTGCACC 480
DB 455 CAGGATGAGAGAGTCTCTCTGAGGCGGTGATTTCCCTGACAGAGAGTCCCGGAGTGCACC 514
QY 481 CTGGGCTGTGCTTGTGCCCCCTCTGAGTGACACACCAGGTGTACTTAGATGAGAGAGGGGGC 540
DB 515 CTGGGCTGTGCTTGTGCCCCCTCTGAGTGACACACCAGGTGTACTTAGATGAGAGAGGGGGC 574
QY 541 TTCAGCGTGACGTCTGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
DB 575 TTCAGCGTGACGTCTGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 634
QY 601 TGGGCCACACTCCAGGATATTGACCAAGCATGTGAGGAGCTCTAGGAGAGGGGCTTGTGTA 660
DB 635 TGGGCCACACTCCAGGATATTGACCAAGCATGTGAGGAGCTCTAGGAGAGGGGCTTGTGTA 694
QY 661 CCGGGTGACAGTCCCTCTACCTGAGCCAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 695 CCGGGTGACAGTCCCTCTACCTGAGCCAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAG 754
QY 721 AGCTGCTCAATGATGAGTGAAGCGGTATGAGCCGAGCTGTCTGCGGCTCCAGCGGCC 780
DB 755 AGCTGCTCAATGATGAGTGAAGCGGTATGAGCCGAGCTGTCTGCGGCTCCAGCGGCC 814
QY 781 GAGCTGCGGGGCTCTCAGAAACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 815 GAGCTGCGGGGCTCTCAGAAACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 874
QY 841 AAAGTGTGATGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 900
DB 875 AAAGTGTGATGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 934
QY 901 GAGCTGCGGGGCTCTCAGAAACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 935 GAGCTGCGGGGCTCTCAGAAACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 994
QY 961 CTGCTGTGAGCAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 995 CTGCTGTGAGCAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1054
QY 1021 GAGTGGAACGAGCAAACTGAGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1055 GAGTGGAACGAGCAAACTGAGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1114
QY 1081 ATGCCCCGGAGATTGACAACTTCACTGAGCGCTTCACTTACCAAAATGTGCGCTC 1140
DB 1115 ATGCCCCGGAGATTGACAACTTCACTGAGCGCTTCACTTACCAAAATGTGCGCTC 1174
QY 1141 TGGATGAGAGATGCGCCAGCTGCTGCGCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1175 TGGATGAGAGATGCGCCAGCTGCTGCGCACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1234
QY 1201 GCTGCAAGAGACAGAGGACCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260

D	1235	GCTGCAAGAGCACAGGGGCAACCCACGTCGTGTTCCACTGTGCAAGATGGGCGTCAGCCGCTCA	1294
Q	1261	GCGGCCACAGTGTGCTGGCCTATGCCA TGAAGCAGTACGAATGCAGCCTGGAGCAGGCGCTG	1320
D	1295	GCGGCCACAGTGTGCTGGCCTATGCCA TGAAGCAGTACGAATGCAGCCTGGAGCAGGCGCTG	1354
Q	1321	CGCCACGTGCAGGAGCTCCGGGCCATCGCCCGCCCAACCTTGCTTCTGCGCCAGCTG	1380
D	1355	CGCCACGTGCAGGAGCTCCGGGCCATCGCCCGCCCAACCTTGCTTCTGCGCCAGCTG	1414
Q	1381	CAGATCTACCAGGGCATCTGTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCAGAAAGTG	1440
D	1415	CAGATCTACCAGGGCATCTGTGACGGCCAGCCGCCAGAGCCATGTCTGGAGCAGAAAGTG	1438
Q	1441	GGTGGGGTCTCCCCAGAGAGCAACCCGCTGAAGTCTACACCATTTCCACCTCTT	1500
D	1439	-----	1438
Q	1501	CCGCCAGAACCTGAGGGTGTGGGGAGAGAGAAGTTGTAGGCATGGAAGAGAGCCAGGCA	1560
D	1439	--GCCAGAACCTGAGGGTGTGGGGAGAGAGAAGTTGTAGGCATGGAAGAGAGCCAGGCA	1496
Q	1561	GCCCCGAAAGAGAGCCTGGGCCACGGCCACGTATAAACCCTCCGAGGGGTATAGGTCC	1620
D	1497	GCCCCGAAAGAGAGCCTGGGCCACGGCCACGTATAAACCCTCCGAGGGGTATAGGTCC	1556
Q	1621	ATCAGTCTTCTGAGACCTCTTGAGCTTGAGAGCACTCAGAGACCAGTGAATGCCA	1680
D	1557	ATCAGTCTTCTGAGACCTCTTGAGCTTGAGAGCACTCAGAGACCAGTGAATGCCA	1616
Q	1681	GAGGTCTTCTTCCACAGAGTCTTACATGAAGACCTTGACGCCCTTCCACAGCTT	1740
D	1617	GAGGTCTTCTTCCACAGAGTCTTACATGAAGACCTTGACGCCCTTCCACAGCTT	1676
Q	1741	GCAAGAGCAAGGAGGAGCCAGCAGGTGACAGGGGGCTCAGCCTGCGCTGAAGTCCGC	1800
D	1677	GCAAGAGCAAGGAGGAGCCAGCAGGTGACAGGGGGCTCAGCCTGCGCTGAAGTCCGC	1736
Q	1801	CAGTCAGTGTATACCTCCAGGGCAGTGCCTGTGTGCCAACCGGACCCAGGCTTCCAG	1860
D	1737	CAGTCAGTGTATACCTCCAGGGCAGTGCCTGTGTGCCAACCGGACCCAGGCTTCCAG	1796
Q	1861	GAGCAGAGCAGGGGCGAGGGGCGAGGGGAGAGGCCCTGCATTTCTCTACGCCCAGG	1920
D	1797	GAGCAGAGCAGGGGCGAGGGGCGAGGGGAGAGGCCCTGCATTTCTCTACGCCCAGG	1856
Q	1921	TTCCGGAAGGTGCTGAGACAGGCCACGCTGCATGACAGTGGAGAGGAGGGCGAGGCTTGA	1980
D	1857	TTCCGGAAGGTGCTGAGACAGGCCACGCTGCATGACAGTGGAGAGGAGGGCGAGGCTTGA	1916

RESULT	5
ABQ73249	
ID	ABQ73249 standard; cDNA; 2704 BP.
XX	
AC	ABQ73249;
XX	
DT	30-SEP-2002 (first entry)
XX	
DE	Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:1
XX	
KW	Human; phosphatase; mitogen activated protein kinase phosphatase;
KW	MAP kinase; enzyme; chromosome 11; gene; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	5' UTR
FT	
FT	CDS
FT	
FT	3' UTR

[illegible]

Db	514	CAGGATGAGACGGTCTCTCTCTGGGGCGGTGGATTTCCTTGACAGCAGCTCCCCCAGCTGCACC	573
QY	481	CTGGGCTGTCTTGGCCCTCTGGAGTGACACCCAGGTGTACTTAGATGAGACGGGGCC	540
Db	574	CTGGGCTGTCTTGGCCCTCTGGAGTGACACCCAGGTGTACTTAGATGAGACGGGGCC	633
QY	541	TTACGCGTGACGCTGTGGGGCAAAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG	600
Db	634	TTACGCGTGACGCTGTGGGGCAAAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG	693
QY	601	TGGGCCACACTCCAGGTATTGCACCAGCATGTGAGCAGCTCTAGGACGGCCTTGT	660
Db	694	TGGGCCACACTCCAGGTATTGCACCAGCATGTGAGCAGCTCTAGGACGGCCTTGT	753
QY	661	CCGGGTGCGAGTGCCCTCACTGGGCCAGCCACTACAGAGAGACTGAACTCCGAACAG	720
Db	754	CCGGGTGCGAGTGCCCTCACTGGGCCAGCCACTACAGAGAGACTGAACTCCGAACAG	813
QY	721	AGCTGCCTCAATGAGTGGACGGCTATGGCCGACCTGGAGTCTCTGGGCTCCAGGCC	780
Db	814	AGCTGCCTCAATGAGTGGACGGCTATGGCCGACCTGGAGTCTCTGGGCTCCAGGCC	873
QY	781	GAGCCTGGCGGGTCTCTAGAACAGAGACAGATGAGACAGCGATCCGTGCTGAGCTGG	840
Db	874	GAGCCTGGCGGGTCTCTAGAACAGAGACAGATGAGACAGCGATCCGTGCTGAGCTGG	933
QY	841	AAAGTGTGGATGTCACTGACCTGGAGAGTGTCACTTCCAAAGATCCGCCAGGCTCTG	900
Db	934	AAAGTGTGGATGTCACTGACCTGGAGAGTGTCACTTCCAAAGATCCGCCAGGCTCTG	993
QY	901	GAGCTGCGCCTGGGGCTCCCCCTCCAGCAGTACCCTGACTTCACTGACACACAGATGCTG	960
Db	994	GAGCTGCGCCTGGGGCTCCCCCTCCAGCAGTACCCTGACTTCACTGACACACAGATGCTG	1053
QY	961	CTGCTGTGTGCGACAGCGGGACCGAGCCTCCCGCATTTTCCCCCACTCTACCTGGGCTCA	1020
Db	1054	CTGCTGTGTGCGACAGCGGGACCGAGCCTCCCGCATTTTCCCCCACTCTACCTGGGCTCA	1113
QY	1021	GAGTGGAAACGACGAAACCTTGGAGGAGCTGCAGAGGAACAGGGTCAACCAATCTTTGAA	1080
Db	1114	GAGTGGAAACGACGAAACCTTGGAGGAGCTGCAGAGGAACAGGGTCAACCAATCTTTGAA	1173
QY	1081	ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATATGCGCCTC	1140
Db	1174	ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATATGCGCCTC	1233
QY	1141	TGGGATGAGAGTGGCCCACTGCTGCCGACTGGAAGAGACGACCGCTTCAATTGAG	1200
Db	1234	TGGGATGAGAGTGGCCCACTGCTGCCGACTGGAAGAGACGACCGCTTCAATTGAG	1293
QY	1201	GCTGCAAGAGCACAGGGCAACCAAGTGTGCTGTCCACTGCAAGATGGGCGTCAGCCGCTCA	1260
Db	1294	GCTGCAAGAGCACAGGGCAACCAAGTGTGCTGTCCACTGCAAGATGGGCGTCAGCCGCTCA	1353
QY	1261	GCGGCCACAGTGTGGCCTATGCCATGAAGCAGTACGAATGCAGCCTGGAGCAGGCCCTG	1320
Db	1354	GCGGCCACAGTGTGGCCTATGCCATGAAGCAGTACGAATGCAGCCTGGAGCAGGCCCTG	1413
QY	1321	CGCCACGTGCAGGAGCTCCGGGCCATGCCCCGCCCAACCTTGCTTCTGCGCAGCTG	1380
Db	1414	CGCCACGTGCAGGAGCTCCGGGCCATGCCCCGCCCAACCTTGCTTCTGCGCAGCTG	1473
QY	1381	CAGATCTACAGGGGATCTCTGACGGCCAGCCGCCAGAGCCATGTCTGGGAGCAGAAAGTG	1440
Db	1474	CAGATCTACAGGGGATCTCTGACGGCCAGCCGCCAGAGCCATGTCTGGGAGCAGAAAGTG	1497
QY	1441	GGTGGGGTCTCCCCAGAGGAGCACCCAGCCCTGAAGTCTTACACCATTCCACCTCTT	1500
Db	1498	-----	1497
QY	1501	CCGCCAGAACCTGAGGGTGTGGGGAGGAGAGAGTTGTAGGCATGGAAGAGGCCAGGCA	1560
Db	1498	CCGCCAGAACCTGAGGGTGTGGGGAGGAGAGAGTTGTAGGCATGGAAGAGGCCAGGCA	1555

QY	1561	GCCCCGAAAGAAAGAGCTGGGCCACGGCCACGTATAAACCTCCGAGGGGTCAAGAGTCC	1620
Db	1556	GCCCCGAAAGAAAGAGCTGGGCCACGGCCACGTATAAACCTCCGAGGGGTCAAGAGTCC	1615
QY	1621	ATCAGTCTTCTGAGAGCCCTCCTTGAGCTGAGAGCACTCAGAGACCAGTGAATGCCA	1680
Db	1616	ATCAGTCTTCTGAGAGCCCTCCTTGAGCTGAGAGCACTCAGAGACCAGTGAATGCCA	1675
QY	1681	GAGGTCTTCTTCCACGAGTCTTCACATGAAGAGCTCTGCAGCCCTTCCACAGCTT	1740
Db	1676	GAGGTCTTCTTCCACGAGTCTTCACATGAAGAGCTCTGCAGCCCTTCCACAGCTT	1735
QY	1741	GCAAGAGCCAAAGGAGAGCCAGCAGGTGACAGAGGGGCTCAGCCTGCCTGAAGTCCCGC	1800
Db	1736	GCAAGAGCCAAAGGAGAGCCAGCAGGTGACAGAGGGGCTCAGCCTGCCTGAAGTCCCGC	1795
QY	1801	CAGTCAGTGTACCTCCAGGGCAGTGCCGTGTGGCCAACCGGACCCAGGCTTCCAG	1860
Db	1796	CAGTCAGTGTACCTCCAGGGCAGTGCCGTGTGGCCAACCGGACCCAGGCTTCCAG	1855
QY	1861	GAGCAGGAGCAGGGGCAAGGGGCAAGGGGCAAGGAGAGCCCTGCATTCTCTACGCCAGG	1920
Db	1856	GAGCAGGAGCAGGGGCAAGGGGCAAGGGGCAAGGAGAGCCCTGCATTCTCTACGCCAGG	1915
QY	1921	TTCCGGAAGGTGGTGAACAGGCCAGCGTGCATGAACAGTGAGAGAGAGGCGAGGCTTGA	1980
Db	1916	TTCCGGAAGGTGGTGAACAGGCCAGCGTGCATGAACAGTGAGAGAGAGGCGAGGCTTGA	1975

RESULT 6

ID ABQ73252 standard; cDNA; 2704 BP

AC ABQ73252;

DT 30-SEP-2002 (first entry)

Human MAP kinase phosphatase splice form 1 cDNA sequence SEQ ID NO:7.

Human; phosphatase; mitogen activated protein kinase phosphatase; MAP kinase; enzyme; chromosome 11; single nucleotide polymorphism; KM

KW SNP; gene; ss.

OS Homo sapiens.

FH Key

ET :

ET :

FT variation

333

VALUATION
EFFECT

ET 3' UTR

variation

三

YY WO200242436-2
PN

XX
PD 30-MAY-2002

XX
DE
07-NOV-2001:

30 NOV - 2000.

PR 18-JAN-2001;
xx

```

Location/Qualifiers
1..93
/*tag= a
94..1509
/*tag= b
/product= "MAP kinase phosphatase splice form 1"
replace(577,A)
/*tag= c
/*standard_name= "single nucleotide polymorphism (SNP)"
replace(1451,G)
/*tag= d
/*standard_name= "single nucleotide polymorphism (SNP)"
1510..2704
/*tag= e
replace(2641,A)
/*tag= f
/*standard_name= "single nucleotide polymorphism (SNP)"

```

01WO-US42995.

00US-0715177.

C
H
C
E

C
-
C
H
C
.
.

PA (PEKE) PE CORP NY.
XX
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-575237/61.
DR P-PSDB; ABP51653.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
PS Claim 1; Fig 3A; 85pp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 1 from the present
CC invention.
XX
SQ Sequence 2704 BP; 569 A; 874 C; 794 G; 467 T; 0 other;

Query Match 89.6%; Score 1774; DB 24; Length 2704;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1882; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

QY 1 ATGGCCCTGTGTCACAGTGAACCGGTTGCCCCCGGGGAGCGCGCTCCACGCCCCGTGGGG 60
Db 94 ATGGCCCTGTGTCACAGTGAACCGGTTGCCCCCGGGGAGCGCGCTCCACGCCCCGTGGGG 153
QY 61 CCCTGGGACCAAGCGGCTCCAGGAGAGTGAAGTCTCCAGGAGGAGAGAGCTTTGGGTTG 120
Db 154 CCCTGGGACCAAGCGGCTCCAGGAGAGTGAAGTCTCCAGGAGGAGAGAGCTTTGGGTTG 213
QY 121 CTCCGTGGGGGTGTCTCTGGGACTGACAGATGAGGGGACAATGATGATGACAGAGAGCC 180
Db 214 CTCCGTGGGGGTGTCTCTGGGACTGACAGATGAGGGGACAATGATGATGACAGAGAGCC 273
QY 181 AGTTCTGAGCCCAAGAGAGAGGCGCCGAGTGAAGAGAGTCCACGGGGACACAGAGAGC 240
Db 274 AGTTCTGAGCCCAAGAGAGAGGCGCCGAGTGAAGAGAGTCCACGGGGACACAGAGAGC 333
QY 241 TTGGGGCAAGGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 334 TTGGGGCAAGGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
QY 301 ATGTACAGCTGTGAGGCGGAGAGATGATCCGCTGGCAGGCCAGCTGAGGCAACC 360
Db 394 ATGTACAGCTGTGAGGCGGAGAGATGATCCGCTGGCAGGCCAGCTGAGGCAACC 453
QY 361 CGGCTCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 454 CGGCTCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAGAGAGAGAGAGAGAGAG 513
QY 421 CAGGATGAGACGGTCTCTGGGCGTGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAG 480
Db 514 CAGGATGAGACGGTCTCTGGGCGTGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAG 573
QY 481 CTGGGCGTGTCTGCCCCCTGTGAGTGAACCCAGAGTGTACTTAGATGAGAGCGGGGGC 540
Db 574 CTGGGCGTGTCTGCCCCCTGTGAGTGAACCCAGAGTGTACTTAGATGAGAGCGGGGGC 633
QY 541 TTCAGCGTGAAGTGTGTGGGCAAGCCGAGATCTTCAAGCCCATCTCCATCCAGAGCATG 600
Db 634 TTCAGCGTGAAGTGTGTGGGCAAGCCGAGATCTTCAAGCCCATCTCCATCCAGAGCATG 693
QY 601 TGGGCCACACTCCAGGTATTGCAACCAAGATGTGAGGAGAGCTTGAAGCAGCGGCTTGTG 660

Db 694 TGGGCCACACTCCAGGTATTGCAACCAAGATGTGAGGAGAGCTTGAAGCAGCGGCTTGTG 753
QY 661 CCGGCTGGAGTGGCCCTCACCTGGGGCCAGCCACTTACAGAGAGAGAGAGAGAGAGAG 720
Db 754 CCGGCTGGAGTGGCCCTCACCTGGGGCCAGCCACTTACAGAGAGAGAGAGAGAGAGAG 813
QY 721 AGCTGCTCAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 814 AGCTGCTCAATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 873
QY 781 GAGCTGGGCGGCTCTCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 874 GAGCTGGGCGGCTCTCAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
QY 841 AAAGTGTGGATGTCAAGTGAAGTGAAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 900
Db 934 AAAGTGTGGATGTCAAGTGAAGTGAAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 993
QY 901 GAGTGGCGCTGGGGCTCCCGCTCCAGCAGTACCGTGAATCTTCAATGAGAGAGAGAGAG 960
Db 994 GAGTGGCGCTGGGGCTCCCGCTCCAGCAGTACCGTGAATCTTCAATGAGAGAGAGAGAG 1053
QY 961 CTGCTGTGGCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 1054 CTGCTGTGGCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1113
QY 1021 GAGTGAACGACGAAACCTGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1114 GAGTGAACGACGAAACCTGAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173
QY 1081 ATGGCCCGGAGATGACAACTTCTACCTGAGCGCTTACCTACCAATATGCGGCTC 1140
Db 1174 ATGGCCCGGAGATGACAACTTCTACCTGAGCGCTTACCTACCAATATGCGGCTC 1233
QY 1141 TGGGATGAGAGAGTGGGCCAGCTGTGCGGACTGGAAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1234 TGGGATGAGAGAGTGGGCCAGCTGTGCGGACTGGAAGAGAGAGAGAGAGAGAGAGAG 1293
QY 1201 GCTGCAAGAGACACAGGGGACCCACAGTGTGTGTCTGCAAGATGGGCGTCAAGCCGCTCA 1260
Db 1294 GCTGCAAGAGACACAGGGGACCCACAGTGTGTGTCTGCAAGATGGGCGTCAAGCCGCTCA 1353
QY 1261 GCGGCCACAGTGTGCTGGCTATGCCATGAAGAGAGTGAATGACAGCTTGAAGAGAGAG 1320
Db 1354 GCGGCCACAGTGTGCTGGCTATGCCATGAAGAGAGTGAATGACAGCTTGAAGAGAGAG 1413
QY 1321 CGCCAGTGAAGAGAGTCCGGCCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1380
Db 1414 CGCCAGTGAAGAGAGTCCGGCCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1473
QY 1381 CAGATCTACAGAGGAGATCTGACGGCCAGCGCCAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1474 CAGATCTACAGAGGAGATCTGACGGCCAGCGCCAGAGAGAGAGAGAGAGAGAGAGAG 1497
QY 1441 GGTGGGCTTCCCAAGAGAGACCCAGCCCTGAAGTCTTACACCATTCCCACTCTT 1500
Db 1498 ----- 1497
QY 1501 CCGCCAGAACCTGAGGCTGTGGGAGAGAGAGTGTGAGGAGAGAGAGAGAGAGAGAGAG 1560
Db 1498 --GCCAGAACCTGAGGCTGTGGGAGAGAGAGTGTGAGGAGAGAGAGAGAGAGAGAG 1555
QY 1561 GCGCCGAAAGAGAGAGCTGGGCCAGGCGCAAGTATAAACCTCCGAGGGGTATGAGGTCC 1620
Db 1556 GCGCCGAAAGAGAGAGCTGGGCCAGGCGCAAGTATAAACCTCCGAGGGGTATGAGGTCC 1615
QY 1621 ATCAGTCTTCTGAGGCTCTCTTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1616 ATCAGTCTTCTGAGGCTCTCTTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1675
QY 1681 GAGGTCTTCTTCCAGAGAGTCTTCAATGAAGAGAGCTTGAAGCCCTTCCACAGCTT 1740

Db	1676	GAGGTCCTTCTCTTCCACGAGTCCTCACATGAAGAGCCTCTGCAGCCCTTCCACAGCTT	1735
QY	1741	GCAAGGACCAAGGGAGGAGGCCACAGGTGGACAGGGGGCTCAGCCTGGCCCTGAAGTCCGC	1800
Db	1736	GCAAGGACCAAGGGAGGAGGCCACAGGTGGACAGGGGGCTCAGCCTGGCCCTGAAGTCCGC	1795
QY	1801	CAGTCAGTGTATACCCCTCCAGGGCAGTGCCTGGTGGCCAAACCGACCCAGGCCCTTCAG	1860
Db	1796	CAGTCAGTGTATACCCCTCCAGGGCAGTGCCTGGTGGCCAAACCGACCCAGGCCCTTCAG	1855
QY	1861	GAGCAGGACGAGGGGCGAGGGGCGAGGGCAGGGAGAGCCCTGCATTTCCTCTACGCCCAGG	1920
Db	1856	GAGCAGGACGAGGGGCGAGGGGCGAGGGGCGAGGGAGAGCCCTGCATTTCCTCTACGCCCAGG	1915
QY	1921	TTCCGGAAAGGTGGTGAAGACAGGCCAGCGTGCAATGACAGTGGAGAGAGGGCGAGGCTGA	1980
Db	1916	TTCCGGAAAGGTGGTGAAGACAGGCCAGCGTGCAATGACAGTGGAGAGAGGGCGAGGCTGA	1975

RESULT 7
AB140805
ID AB140805 standard; cDNA; 2322 BP.

AC ABL40805;

DT 03-JUL-2002 (first entry)

DE Human MAP kinase phosphatase-like enzyme encoding cDNA

KM Mitogen activated protein, MAP; MAP kinase phosphatase-like enzyme;
KM antiasthmatic; antidiabetic; anorectic; cyostatic; cardant; human;
KM antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KM neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KM antiallergic; dermatological; vulnery; gene therapy; ss.

OS Homo sapiens.

FH	Key	Location/Qualifiers
FT	CDS	1..2322

```

FT      /product= "MAP kinase phosphatase-like enzyme"
FT      /note= "contains internal codon deletions"
FT      /transl_except= "(pos: 156..157, aa: Leu)"
FT      /note= "there is an apparent one codon deletion which
FT      alters the reading frame"
FT      /transl_except= "(pos: 180..181, aa: Arg)"
FT      /note= "there is an apparent one codon deletion which
FT      alters the reading frame"
FT      /transl_except= "(pos: 414..415, aa: Val)"
FT      /note= "there is an apparent one codon deletion which
FT      alters the reading frame"
FT      /transl_except= "(pos: 1197..1198, aa: Tyr)"
FT      /note= "there is an apparent one codon deletion which
FT      alters the reading frame"
FT      /transl_except= "(pos: 2088..2089, aa: Xaa)"
FT      /note= "there is an apparent one codon deletion which
FT      alters the reading frame"

```

PN WO200220732-A2.

PD 14-MAR-2002.

PF 27-AUG-2001; 2001WO-EP09848.

PR 07-SEP-2000; 2000US-230709P.

PA (FARB.) BAYER AG.

PI Liou J;

DR WPI; 2002-339802/37

XX

PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases -
XX
PS Claim 1, Fig 13; 134pp; English.

PS Claim 1; Fig 13; 134pp; English.

CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischaemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme polypeptide encoding cDNA.

Sequence 2322 BP; 467 A; 751 C; 718 G; 386 T; 0 other;

Query Match	85.1%;	Score 1684;	DB 24;	Length 2322;
Best Local Similarity	95.0%;	Pred. No. 0;		
Matches 1852; Conservative	0;	Mismatches 75;	Indels 23;	Gaps 10

39 CGCGCCTCCACGCCCGTGGGCCCTGGGACCAAGCGGTCCAGCGAAGGAGTGGACTCCA 98

Db 147 CGCCCTCACCTGGGCTGCTCTCTCGGCAAGGACCTGGTCCAGCGAAGGAGTCCGACTCCA 200

99 GCGAAGCAGAGCTTTGCGGTGCTCC-GTGGGGCTGTCTCTGGACTGCAGGATGGAGGG 150

Db 207 GCGA---AGAGCTTGGCGTGCTCCTGTGGGCTGTCTCTGGACTGCAGGATGGAGGGC 26

158 ACAATGATGTCAGCAGAGGCCAGTTCCTGAGCCACACAGAGAAGGCCCCCGAGTGAGGAGC 21

Db 263 ACAATGATGCAGCAGAGGCCAGTCTGAGCCACACAGAGAAGCCCCGAGTGAGGAGC 32

218 AGCTCCACGGG--ACCAAGACAGACTT--CGGCAAGATCCCAAGTCCCCAGAAACA 27

Db 323 AGCTCCACGGGGAACCCAGACAGACTTCGGTGCAGGATCCAGAGTCCCAGAGCA 38

273 GGAGGAGCAGAGCGCAGCACTGCACCTCATGTGTACAGCTGCTGAGGCCCGCAGGATGACAT 33

Db 383 GGAGGAGCAGAGGCACTGCAACCTCATGAGCGTGCTGAGGCCGACGGATGACAT 44

333 CCGCCTGGCA--GCCCAGCTGGAGGCACCCCGGCTCCCGGCTCCGCTACCTGC----- 38

Db 443 CCGACTTGGAAGCCCCAAGCTGGAGGCCACCCCGGGCTCCCGGATCCGGATACCTTGCT 50

386 TGGTAGTTTCTACACGAGAA-CGAGGAAGTCTGAG-CCAGGATGAGACCGTCTCTGCGG 44

db 503 TGGTACTTCTACACGAGAGGAGAGTCTGAGCCCGAGATGAGACCGTCTCTGGG 56

444 -CGTGGATTTCCTGACAGCAGCTCCCCCAGCTGCACCCCTGGGCTTGGTCTTGGCCCCCTCT 50

Db 563 ACGTGATTccccTGACGACGCTccccGAGCTGACccctGGGccTGGTCTTGGccctCT 62

503 GGAAGTGAACACCCAGGTGTACTTAGATGGAGACGGGGGGCTTCAGCGTGAAGTCTGTGGGC 56

623 GGAAGTGAACACCAAGTGTACTTATATGAGACCGGGGGCTTCAGCCGTGACGTCTGTGGGC 68

563 AAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGTTATTGC 62

Db 683 AAACCCGGATCTTCAAGCCCATCTCCATCCAGACCATGTGGGCCACACTCCAGGTATTGC 742
QY 623 ACCAAGCATGTGAGGACAGCTCTAGGCAGCGGCTTGTACCGGGGTGGAGTGCCTCACCCT 682
Db 743 ACCAAGCATGTGAGGACAGCTCTAGGCAGCGGCTTGTACCGGGGTGGAGTGCCTCACCCT 802
QY 683 GGGCCAGCCACTACCAAGAGAGACTGAATCCGAACAGAGCTGCTCAATGAGTGAACGG 742
Db 803 GGGCCAGCCACTACCAAGAGAGACTGAATCCGAACAGAGCTGCTCAATGAGTGAACGG 862
QY 743 CTATGCGCCGACCTGGAGTCTCTGCGGCTCCAGCGCCGAGCCTGCGGGTCTCAGAAC 802
Db 863 CTATGCGCCGACCTGGAGTCTCTGCGGCTCCAGCGCCGAGCCTGCGGGTCTCAGAAC 922
QY 803 AGGACGAGATGAGCAGCGCGATCCGTGCTGAGCTGTGAAAGTGTGATGTCAAGTACC 862
Db 923 AGGACGAGATGAGCAGCGCGCGATCCGTGCTGAGCTGTGAAAGTGTGATGTCAAGTACC 982
QY 863 TGGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTGAGCTGCGGCTGGGGCTCCCC 922
Db 983 TGGAGAGTGTGCTTCCAGAGAGATCCGCCAGGCTCTGAGCTGCGGCTGGGGCTCCCC 1042
QY 923 TCCAGCAGTACCGTGACTTTCATCGACACACAGATGCTGCTGTGTGGACACAGCGGACC 982
Db 1043 TCCAGCAGTACCGTGACTTTCATCGACACACAGATGCTGCTGTGTGGACACAGCGGACC 1102
QY 983 GAGCTTCCCGCATCTTCCCGCACTTACCTGGGCTCAGAGTGAACCGCAACCACTGG 1042
Db 1103 GAGCTTCCCGCATCTTCCCGCACTTACCTGGGCTCAGAGTGAACCGCAACCACTGG 1162
QY 1043 AGGAGCTGCAGAGGAACAGGGGTCAACCACATTTGAACATGGCCCGGAGATTGACAACT 1102
Db 1163 AGGAGCTGCAGAGGAACAGGGGTCAACCACATTTG--ATGGCCCGGAGATTGACAACT 1219
QY 1103 TCTACCTTGAGCGGCTTACCTACCAACAATGTGCGCTCTGGATGAGAGTGGCCACAGC 1162
Db 1220 TCTACCTTGAGCGGCTTACCTACCAACAATGTGCGCTCTGGATGAGAGTGGCCACAGC 1279
QY 1163 TGTGCGCGCACTGGAAGAGAGCAACCGCTTCAATTGAGGCTGCAGAGCAAGGGCACACC 1222
Db 1280 TGTGCGCGCACTGGAAGAGAGCAACCGCTTCAATTGAGGCTGCAGAGCAAGGGCACACC 1339
QY 1223 ACGTGTGTGTCACCTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGCTATG 1282
Db 1340 ACGTGTGTGTCACCTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGCTATG 1399
QY 1283 CCATGAAGCAGTACGAATGAGAGCTGAGCAGGCGCTGCGCACAGTGAAGAGCTCCGGC 1342
Db 1400 CCATGAAGCAGTACGAATGAGAGCTGAGCAGGCGCTGCGCACAGTGAAGAGCTCCGGC 1459
QY 1343 CCATGCGCCGCGCCCAACCTGCGCTTCTGCGCGCAGCTGCAGATCTACAGAGGCATCTGA 1402
Db 1460 CCATGCGCCGCGCCCAACCTGCGCTTCTGCGCGCAGCTGCAGATCTACAGAGGCATCTGA 1519
QY 1403 CGGCCAGCGCCAGAGCCATGTTGGAGCAGAAAGTGGGTGGGTCTCCCCAGAGAGAGC 1462
Db 1520 CGGCCAGCGCCAGAGCCATGTTGGAGCAGAAAGTGGGTGGGTCTCCCCAGAGAGAGC 1579
QY 1463 ACCCAGCCCCCTGAAGTCTCTACACCATTTCCCACTTCTCCGCGCAGAACTGAGGGTGTG 1522
Db 1580 ACCCAGCCCCCTGAAGTCTCTACACCATTTCCCACTTCTCCGCGCAGAACTGAGGGTGTG 1639
QY 1523 GGGAGAGAGAGGTGTGAGGCATGGAAGAGAGCCAGCGCAGCCCCGAAAGAGAGCCTGGGC 1582
Db 1640 GGGAGAGAGAGGTGTGAGGCATGGAAGAGAGCCAGCGCAGCCCCGAAAGAGAGCCTGGGC 1699
QY 1583 CACGGCCACGTATAAACCCTCCAGAGGGGTGATGAGGTCCATCAAGTCTTTCGAGCCCTCCT 1642
Db 1700 CACGGCCACGTATAAACCCTCCAGAGGGGTGATGAGGTCCATCAAGTCTTTCGAGCCCTCCT 1759
QY 1643 TGGAGCTGGAAGACACCTCAAGAGCCAGTGAACATGCCAGAGGTTCTTTCACAGAGT 1702
Db 1760 TGGAGCTGGAAGACACCTCAAGAGCCAGTGAACATGCCAGAGGTTCTTTCACAGAGT 1819

QY 1703 CTTACATGAGAGCCCTTGACAGCCCTTCCACAGCTTGCAAGGACCAAGGAGGCCAGC 1762
Db 1820 CTTACATGAGAGCCCTTGACAGCCCTTCCACAGCTTGCAAGGACCAAGGAGGCCAGC 1879
QY 1763 AGGTGACAGGGGGCTCAGCCTGCCCCCTGAAGTCCCGCAGTCAAGTGTACCTCCAGG 1822
Db 1880 AGGTGACAGGGGGCTCAGCCTGCCCCCTGAAGTCCCGCAGTCAAGTGTACCTCCAGG 1939
QY 1823 GCAGTGCCTGTGGCCACCCGACCCAGGCTTCCAGAGCAGAGCAGGGGCGAGGGGC 1882
Db 1940 GCAGTGCCTGTGGCCACCCGACCCAGGCTTCCAGAGCAGAGCAGGGGCGAGGGGC 1999
QY 1883 AGGGGCAAGGAGAGCCCTGCATTTCTCTACGCCCCAGGTTCCGGAAGGTGTGAGACAGG 1942
Db 2000 AGGGGCAAGGAGAGCCCTGCATTTCTCTACGCCCCAGGTTCCGGAAGGTGTGAGACAGG 2059
QY 1943 CCAGCGTGATGACAGTGAAGAGAGGGCG 1972
Db 2060 CCAGCGTGATGACAGTGAAGAGAGGGCG 2089

RESULT 8
ABQ73251
ID ABQ73251 standard; cDNA; 2540 BP.
XX
AC ABQ73251;
DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 3 cDNA sequence SEQ ID NO:3.
XX
KW Human; phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11; gene; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..87
FT CDS /*tag= a
FT 88..1314
FT /*tag= b
FT 3'UTR /product= "MAP kinase phosphatase splice form 3"
FT 1315..2540
FT /*tag= c
XX
PN WO200242436-A2.
XX
PD 30-MAY-2002.
XX
PE 07-NOV-2001; 2001WO-US42995.
XX
PR 20-NOV-2000; 2000US-0715177.
PR 18-JAN-2001; 2001US-0761640.
XX
PA (PEKE) PE CORP NY.
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX WPI; 2002-575237/61.
DR P-PSDB; ABP51655.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies -
XX
PS Claim 1; Fig 1C; 85pp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted

CC mixture to determine whether a complex is formed with the agent bound
CC (1). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (1) and the polynucleotide
CC sequences encoding (1) can be used in gene therapy. The present sequence
CC encodes human MAP kinase phosphatase splice form 3 from the present
CC invention.
XX

Sequence 2540 BP; 552 A; 809 C; 739 G; 440 T; 0 other;

Query Match 68.0%; Score 1345.6; DB 24; Length 2540;
Best Local Similarity 85.1%; Pred. No. 3e-275;
Matches 1689; Conservative 0; Mismatches 4; Indels 291; Gaps 5;

QY 1 ATGCCCCCTGGTCAAGTGAAGCCGTTGCCCCCGGCGACGGCGCCCTCCACGCCCTGGGG 60
DB 88 ATGGCCCTGTGTCAGTGAAGCCGTTGCCCCCGGCGACGGCGCCCTCCACGCCCTGGGG 147
QY 61 CCCTGGACCAAGCGCGTCCAGCGAAGAGTCACTCCAGCGAAGCAGAGCTTTGCGGTG 120
DB 148 CCCTGGACCAAGCGCGTCCAGCGAAGAGTCACTCCAGCGAAGCAGAGCTTTGCGGTG 207
QY 121 CTCCGTGGGGTGTCTCTGGAGTGAAGATGAGGGGACATGATGATGACAGAGGCC 180
DB 208 CTCCGTGGGGTGTCTCTGGAGTGAAGATGAGGGGACATGATGATGACAGAGGCC 267
QY 181 AGTTCTGAGCCACAGAGAGAGGCCCCGAGTGAAGAGAGTCCACGGGACCAAGACAGAC 240
DB 268 AGTTCTGAGCCACAGAGAGAGGCCCCGAGTGAAGAGAGTCCACGGGACCAAGACAGAC 327
QY 241 TTCGGGCAAGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 328 TTCGGGCAAGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
QY 301 ATGTACAGCTGTGAGGGCCGAGAGATGACATCCGCTGGCAGCCCAAGCTGAGGACACC 360
DB 388 ATGTACAGCTGTGAGGGCCGAGAGATGACATCCGCTGGCAGCCCAAGCTGAGGACACC 447
QY 361 CGGCTCCCGGCTCCGCTACCTGCTGTGATTTCTACAGAGAGAGAGAGAGAGAGAGAG 420
DB 448 CGGCTCCCGGCTCCGCTACCTGCTGTGATTTCTACAGAGAGAGAGAGAGAGAGAGAG 507
QY 421 CAGGATGAGACGGTCTCTCTGGGCGTGAATTTCCCTGACAGAGAGAGAGAGAGAGAGAG 480
DB 508 CAGGATGAGACGGTCTCTCTGGGCGTGAATTTCCCTGACAGAGAGAGAGAGAGAGAGAG 567
QY 481 CTGGGCTGTGTTGCCCCCTCTGAGTGAACACCAAGGTGACTTGAAGAGAGAGAGAGAGAG 540
DB 568 CTGGGCTGTGTTGCCCCCTCTGAGTGAACACCAAGGTGACTTGAAGAGAGAGAGAGAGAG 627
QY 541 TTCAGCGTGAAGTGTGAGGCAAGCGGATCTTCAAGGCTCATCTCCATCCAGAGCATG 600
DB 628 TTCAGCGTGAAGTGTGAGGCAAGCGGATCTTCAAGGCTCATCTCCATCCAGAGCATG 687
QY 601 TGGGCACTCCAGGTATGACCAAGCATGTGAGGAGCTTAGGAGCGGCTTGTGA 660
DB 688 T----- 688
QY 661 CCGGGTGCAGTGCCTCACTGGGCGACCACTACAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 689 ----- 689
QY 721 AGCTGCTCAATGAGTGAAGCGGCTATGAGCGGAGCTGTGCGGCTCCAGCGCC 780
DB 689 ----- 689
QY 781 GAGCCTGCGGGTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 689 -----GCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 738
QY 841 AAAGTGTGATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 900
DB 739 AAAGTGTGATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 798

QY 901 GAGCTGCGCTGGGGCTCCCCCTCCAGAGTACCGTGAATCTTATGACAAACCATGCTG 960
DB 799 GAGCTGCGCTGGGGCTCCCCCTCCAGAGTACCGTGAATCTTATGACAAACCATGCTG 858
QY 961 CTGCTGTGGACAGCGGGAGCCGAGCTCCCGCATCTTCCCCCACTTACTGGGCTCA 1020
DB 859 CTGCTGTGGACAGCGGGAGCCGAGCTCCCGCATCTTCCCCCACTTACTGGGCTCA 918
QY 1021 GAGTGAACGAGCAAAACCTGAGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 919 GAGTGAACGAGCAAAACCTGAGAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 978
QY 1081 ATGCCCCGGAGATTGACCACTTCTACCTGAGCGCTTCACTACCAATGTCGCTC 1140
DB 979 ATGCCCCGGAGATTGACCACTTCTACCTGAGCGCTTCACTACCAATGTCGCTC 1038
QY 1141 TGGGATGAGAGTGGCCCCAGCTGTGCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1039 TGGGATGAGAGTGGCCCCAGCTGTGCGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1098
QY 1201 GCTGCAAGAGACAGGGGACCCAGCTGTGCTGCTCACTGCAAGATGGGCTGACCGCTCA 1260
DB 1099 GCTGCAAGAGACAGGGGACCCAGCTGTGCTGCTCACTGCAAGATGGGCTGACCGCTCA 1158
QY 1261 GCGGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1159 GCGGCCACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1218
QY 1321 CGCCACGTGCAAGAGTCCCGGCCATCGCCCCCGCCCAACCTGCTGCTGCTGCTGCTGCTG 1380
DB 1219 CGCCACGTGCAAGAGTCCCGGCCATCGCCCCCGCCCAACCTGCTGCTGCTGCTGCTGCTG 1278
QY 1381 CAGATCTACCAAGGCATCTCTGACGGCCAGCCGCCAGGCCATGTCTGGAGAGAGAAAGTG 1440
DB 1279 CAGATCTACCAAGGCATCTCTGACG----- 1302
QY 1441 GGTGGGCTTCCCCAGAGAGACCCAGCCCTGAAGTCTTACACCATTCACCTCTT 1500
DB 1303 ----- 1302
QY 1501 CCGCAGAACCTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB 1303 --GCCAGAACCTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1360
QY 1561 GCGCCGAAAGAGAGAGCTGGG--CCAGCGCCAGATATACTCCGAGGGGTATGAGGT 1618
DB 1361 GCGCCGAAAGAGAGAGCTGGGCGCAGGGGCAAGTATACTCCGAGGGGTATGAGGT 1420
QY 1619 CCATCAGTCTTGTGAGCCCTCTT--GAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1676
DB 1421 CCATCAGTCTTGTGAGCCCTCTTGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1480
QY 1677 GCCAGAGTCTTCTTCCACAGAGTCTTACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1736
DB 1481 GCCAGAGTCTTCTTCCACAGAGTCTTACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
QY 1737 GCTTGAAGGACCAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1796
DB 1541 GCTTGAAGGACCAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1600
QY 1797 CCGCCAGTCAAGTGTACCCCTCAGAGGAGAGTGCCTGTGAGGAGAGAGAGAGAGAGAGAG 1856
DB 1601 CCGCCAGTCAAGTGTACCCCTCAGAGGAGAGTGCCTGTGAGGAGAGAGAGAGAGAGAGAG 1660
QY 1857 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1916
DB 1661 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1720
QY 1917 CAGGTTCCGGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1976
DB 1721 CAGGTTCCGGAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1780

QY 1977 CTGA 1980
Db 1781 CTGA 1784

RESULT 9

AAF63578 standard, cDNA; 1026 BP.

AAF63578;

11-MAY-2001 (first entry)

Human phosphatase NP_060746_h coding sequence.

Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia; cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia; congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease; Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome; schizophrenia; hamartoma; ss.

Homo sapiens.

WO200112819-A2.

22-FEB-2001.

11-AUG-2000; 2000WO-US22158.

13-AUG-1999; 99US-0149005.

(SUGEN-) SUGEN INC.

Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

WPI; 2001-211226/21.
P-PSDB; AAB73226.

New protein phosphatase polypeptide for diagnosing and treating phosphatase related disorders such as cancer, schizophrenia, cardiac dysfunction and/or vascular disorders -
Claim 1; Fig 4; 138bp; English.

The present invention relates to phosphatase proteins and coding sequences. The present sequence is one such phosphatase coding sequence. Phosphatases are enzymes that catalyze the dephosphorylation of proteins modified by phosphorylation of serine, threonine or tyrosine residues. The phosphatases are useful for treating a variety of diseases: for example cancer e.g. breast, urogenital, prostate, head, neck, lung cancers, synovial sarcoma, renal cell carcinoma, non-small cell lung cancer, hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer, glioblastoma, colorectal cancer and thyroid cancer, pathophysiological hypoxia, cardiac dysfunction and/or vascular disorders, myopathies, congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease, ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan Zonana syndrome, schizophrenia and hamartomas.

Sequence 1026 BP; 217 A; 324 C; 326 G; 159 T; 0 other;

Query Match 51.8%; Score 1026; DB 22; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.1e-207;
Matches 1026; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 955 ATGCTGCTGCTGTGGACAGCGGAGCCGAGCCTCCGCACTTCCCCACCTTACCTG 1014
Db 1 ATGCTGCTGCTGTGGACAGCGGAGCCGAGCCTCCGCACTTCCCCACCTTACCTG 60
QY 1015 GGCTCAGAGTGGAACGACGAACTGTGAGAGCTGACAGAGGAGGTCACCCACATC 1074
Db 61 GGCTCAGAGTGGAACGACGAACTGTGAGAGCTGACAGAGGAGGTCACCCACATC 120

QY 1075 TTGAACATGGCCCGGAGATTGACAACTTACCTTGAGCGCTTACCTACCAATGTG 1134
Db 121 TTGAACATGGCCCGGAGATTGACAACTTACCTTGAGCGCTTACCTACCAATGTG 180
QY 1135 CGCCTTGAGATGAGAGTCCGCCAGCTGTCTGCCCACTGGAAGAGACGACCGCTTC 1194
Db 181 CGCCTTGAGATGAGAGTCCGCCAGCTGTCTGCCCACTGGAAGAGACGACCGCTTC 240
QY 1195 ATTGAGCTGCAAGAGACAGGGCACCCACGCTGTGTCCACTGCAAGATGGCGTCAGC 1254
Db 241 ATTGAGCTGCAAGAGACAGGGCACCCACGCTGTGTCCACTGCAAGATGGCGTCAGC 300
QY 1255 CGCTCAGCGGCCACAGTGTGCTGCTATGCTCATGACAGTACGATGACGCTGAGACAG 1314
Db 301 CGCTCAGCGGCCACAGTGTGCTGCTATGCTCATGACAGTACGATGACGCTGAGACAG 360
QY 1315 GCCCTGCGCCACGTGACAGAGCTCCGCCATCGCCCGCCCAACCTGGCTTCTGCGC 1374
Db 361 GCCCTGCGCCACGTGACAGAGCTCCGCCATCGCCCGCCCAACCTGGCTTCTGCGC 420
QY 1375 CAGCTGACATCTACACAGGCAATCTGACGCGCCAGCGCCACAGACCATGTCTGGAGCAG 1434
Db 421 CAGCTGACATCTACACAGGCAATCTGACGCGCCAGCGCCACAGACCATGTCTGGAGCAG 480
QY 1435 AAAGTGGTGGGTCTCCCGACAGAGACACCCAGCCCTGAAGTCTTACACCATTCOCA 1494
Db 481 AAAGTGGTGGGTCTCCCGACAGAGACACCCAGCCCTGAAGTCTTACACCATTCOCA 540
QY 1495 CCTCTTCGCGCAGAACCTGAGGGTGTGGGAGAGAAAGTTGTAGGATGGAAGAGAGC 1554
Db 541 CCTCTTCGCGCAGAACCTGAGGGTGTGGGAGAGAAAGTTGTAGGATGGAAGAGAGC 600
QY 1555 CAGCAGCCCGGAAAGAGAGCCTGGGCCACGCGCAAGTATAACCTCCGAGGGGTGATG 1614
Db 601 CAGCAGCCCGGAAAGAGAGCCTGGGCCACGCGCAAGTATAACCTCCGAGGGGTGATG 660
QY 1615 AGGTCCATGATCTTCTTGAGAGCCTCTTGAGAGTGAAGACACCTCAGAGACAGTGAC 1674
Db 661 AGGTCCATGATCTTCTTGAGAGCCTCTTGAGAGTGAAGACACCTCAGAGACAGTGAC 720
QY 1675 ATGCCAGAGTCTTCTTCCCAAGAGTCTTCAATGAAGACCTTGACGCCCTTCCCA 1734
Db 721 ATGCCAGAGTCTTCTTCCCAAGAGTCTTCAATGAAGACCTTGACGCCCTTCCCA 780
QY 1735 CAGCTTGCAAGACCAAGGAGGCCAGCAGGTGACAGGGGGCTCAGCCTGCCCTGAAG 1794
Db 781 CAGCTTGCAAGACCAAGGAGGCCAGCAGGTGACAGGGGGCTCAGCCTGCCCTGAAG 840
QY 1795 TCCCGCAGTCAAGTGTACCTTCCAGGAGGAGTCCGCTGTGGCCAAACCGGACGAGCC 1854
Db 841 TCCCGCAGTCAAGTGTACCTTCCAGGAGGAGTCCGCTGTGGCCAAACCGGACGAGCC 900
QY 1855 TTCCAGAGCAGAGCAGAGGGGAGGGGAGGGGAGGGAGAGCCCTGCAATTTCTCTACG 1914
Db 901 TTCCAGAGCAGAGCAGAGGGGAGGGGAGGGGAGGGGAGAGCCCTGCAATTTCTCTACG 960
QY 1915 CCCAGTTCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGGGAG 1974
Db 961 CCCAGTTCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGGGAG 1020
QY 1975 GCCTGA 1980
Db 1021 GCCTGA 1026

RESULT 10

ABNS9832 standard, cDNA; 2061 BP.

ABNS9832;

28-JUN-2002 (first entry)

DE Novel human coding sequence SEQ ID NO: 243.

KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag; gene; ss.

OS Homo sapiens.

PN WO200222660-A2.

PD 21-MAR-2002.

PF 10-SEP-2001; 2001WO-US26015.

PR 11-SEP-2000; 2000US-0659671.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou F, Asundi V, Zhang J, Zhao QA, Ren F,

PI Xue AJ, Yang Y, Wehrman T, Dymnac RT;

DR WPI; 2002-292408/33.

DR P-PSDB; ABB97419.

PT An isolated polynucleotide for treating diseases associated with its encoded polypeptide such as cancer and multiple sclerosis -

PS Claim 1; SEQ ID NO 243; 509pp; English

CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.

Sequence 2061 BP; 415 ~~A~~-672 C; 605 G; 369 T; 0 other; 50

Query Match	43.0%;	Score 851;	DB 24;	Length 2061;
Best Local Similarity	100.0%;	Pred. No. 1.4e-170;		
Matches 851;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	ATGGCCCTGGTCAACAGTGAAGCCGTTCCGCCCCCGGGCAGCGCGCTCCACGCCCTGGGG	60
Db	147	ATGGCCCTGGTCAACAGTGAAGCCGTTCCGCCCCCGGGCAGCGCGCTCCACGCCCTGGGG	206
QY	61	CCCTGGGACCAAGCGGCTCCAGCGAAGAGTCCACTCCAGCGAAGCAGAGCTTTGGCGTG	120
Db	207	CCCTGGGACCAAGCGGCTCCAGCGAAGAGTCCACTCCAGCGAAGCAGAGCTTTGGCGTG	266
QY	121	CTCCGTGGGGCTGTCTCTGGGACTGCAGATGGAAGGGGACAAATGATGATGCAGCAGAGCC	180
Db	267	CTCCGTGGGGCTGTCTCTGGGACTGCAGATGGAAGGGGACAAATGATGATGCAGCAGAGCC	326
QY	181	AGTCTGAGCCAAACAGAGAAGCCCCCGAGTGAAGAGAGCTCCACGGGGACCAAGACAGAC	240
Db	327	AGTCTGAGCCAAACAGAGAAGCCCCCGAGTGAAGAGAGCTCCACGGGGACCAAGACAGAC	386
QY	241	TTGGGGCAAGATCCACAGAGTCCCCCAGAAACAGAGGAGCAGAGGACGACCTGCACCTC	300
Db	387	TTGGGGCAAGATCCACAGAGTCCCCCAGAAACAGAGGAGCAGAGGACGACCTGCACCTC	446
QY	301	ATGGTACAGCTGCTGAGGCGGCAGAGATGACATCCGCGCTGGCAGGCCCACTGGAAGCACCC	360
Db	447	ATGGTACAGCTGCTGAGGCGGCAGAGATGACATCCGCGCTGGCAGGCCCACTGGAAGCACCC	506
QY	361	CGGCTCCCCCGGCTCCGCTACCTGCTGGTAGTTTCTACACGAGAAAGGAGAGGTCTGAGC	420

Accession	Source	Length	Strain	Sex	Age	Organ	Method	Ref
AB014722	GenBank	1755	Human	Male	26	Testis	RT-PCR	1
U00096	GenBank	2938	Human	Male	26	Testis	RT-PCR	2
U00097	GenBank	2938	Human	Male	26	Testis	RT-PCR	3
U00098	GenBank	2938	Human	Male	26	Testis	RT-PCR	4
U00099	GenBank	2938	Human	Male	26	Testis	RT-PCR	5
U00100	GenBank	2938	Human	Male	26	Testis	RT-PCR	6
U00101	GenBank	2938	Human	Male	26	Testis	RT-PCR	7
U00102	GenBank	2938	Human	Male	26	Testis	RT-PCR	8
U00103	GenBank	2938	Human	Male	26	Testis	RT-PCR	9
U00104	GenBank	2938	Human	Male	26	Testis	RT-PCR	10
U00105	GenBank	2938	Human	Male	26	Testis	RT-PCR	11
U00106	GenBank	2938	Human	Male	26	Testis	RT-PCR	12
U00107	GenBank	2938	Human	Male	26	Testis	RT-PCR	13
U00108	GenBank	2938	Human	Male	26	Testis	RT-PCR	14
U00109	GenBank	2938	Human	Male	26	Testis	RT-PCR	15
U00110	GenBank	2938	Human	Male	26	Testis	RT-PCR	16
U00111	GenBank	2938	Human	Male	26	Testis	RT-PCR	17
U00112	GenBank	2938	Human	Male	26	Testis	RT-PCR	18
U00113	GenBank	2938	Human	Male	26	Testis	RT-PCR	19
U00114	GenBank	2938	Human	Male	26	Testis	RT-PCR	20
U00115	GenBank	2938	Human	Male	26	Testis	RT-PCR	21
U00116	GenBank	2938	Human	Male	26	Testis	RT-PCR	22
U00117	GenBank	2938	Human	Male	26	Testis	RT-PCR	23
U00118	GenBank	2938	Human	Male	26	Testis	RT-PCR	24
U00119	GenBank	2938	Human	Male	26	Testis	RT-PCR	25
U00120	GenBank	2938	Human	Male	26	Testis	RT-PCR	26
U00121	GenBank	2938	Human	Male	26	Testis	RT-PCR	27
U00122	GenBank	2938	Human	Male	26	Testis	RT-PCR	28
U00123	GenBank	2938	Human	Male	26	Testis	RT-PCR	29
U00124	GenBank	2938	Human	Male	26	Testis	RT-PCR	30
U00125	GenBank	2938	Human	Male	26	Testis	RT-PCR	31
U00126	GenBank	2938	Human	Male	26	Testis	RT-PCR	32
U00127	GenBank	2938	Human	Male	26	Testis	RT-PCR	33
U00128	GenBank	2938	Human	Male	26	Testis	RT-PCR	34
U00129	GenBank	2938	Human	Male	26	Testis	RT-PCR	35
U00130	GenBank	2938	Human	Male	26	Testis	RT-PCR	36
U00131	GenBank	2938	Human	Male	26	Testis	RT-PCR	37
U00132	GenBank	2938	Human	Male	26	Testis	RT-PCR	38
U00133	GenBank	2938	Human	Male	26	Testis	RT-PCR	39
U00134	GenBank	2938	Human	Male	26	Testis	RT-PCR	40
U00135	GenBank	2938	Human	Male	26	Testis	RT-PCR	41
U00136	GenBank	2938	Human	Male	26	Testis	RT-PCR	42
U00137	GenBank	2938	Human	Male	26	Testis	RT-PCR	43
U00138	GenBank	2938	Human	Male	26	Testis	RT-PCR	44
U00139	GenBank	2938	Human	Male	26	Testis	RT-PCR	45
U00140	GenBank	2938	Human	Male	26	Testis	RT-PCR	46
U00141	GenBank	2938	Human	Male	26	Testis	RT-PCR	47
U00142	GenBank	2938	Human	Male	26	Testis	RT-PCR	

XX Claim 8; SEQ ID 12452; 2537bp + CD ROM; English.
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other;
SQ
Query Match 35.1%; Score 694; DB 22; Length 1755;
Best Local Similarity 89.1%; Pred. No. 2.2e-137;
Matches 802; Conservative 0; Mismatches 0; Indels 98; Gaps 1;
QY 1081 ATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGCGCTC 1140
DB 322 ATGCCCCGGAGATTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGCGCTC 381
QY 1141 TGGGATGAGAGTGGCGCCAGCTGTGCGCAGCTGGAAGAGAGCAACCGCTTCATTGAG 1200
DB 382 TGGGATGAGAGTGGCGCCAGCTGTGCGCAGCTGGAAGAGAGCAACCGCTTCATTGAG 441
QY 1201 GCTGCAAGAGCACAGGGGCAACCCAGCTGTGCTGCTCCACTGCAAGATGGCGCTCAGCCGCTCA 1260
DB 442 GCTGCAAGAGCACAGGGGCAACCCAGCTGTGCTGCTCCACTGCAAGATGGCGCTCAGCCGCTCA 501
QY 1261 GCGGCCACAGTGTGCTGCTATGCTCATGATGAGAGTACGATGAGCGCTTGAGAGCGCCCTG 1320
DB 502 GCGGCCACAGTGTGCTGCTATGCTCATGATGAGAGTACGATGAGCGCTTGAGAGCGCCCTG 561
QY 1321 CGCCACGTGACAGAGCTCCGCGCCATCGCCCGCCCAACCCCTGGCTTCTGCGCCAGCTG 1380
DB 562 CGCCACGTGACAGAGCTCCGCGCCATCGCCCGCCCAACCCCTGGCTTCTGCGCCAGCTG 621
QY 1381 CAGATCTACAGGGGCTCTGACGGCCAGCGCCAGAGCCATGCTGGAGAGCAAGATG 1440
DB 622 CAGATCTACAGGGGCTCTGACG----- 645
QY 1441 GGTGGGGTCTCCCGAGAGGACACCCAGCCCTGAAGTCTTACACCATTTCCACCTCTT 1500
DB 646 ----- 645
QY 1501 CCGCCAGAACCTGAGGGTGTGGGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGGCA 1560
DB 646 --GCCAGAACCTGAGGGTGTGGGAGAGAGGTTGTAGGCATGGAAGAGAGCCAGGCA 703
QY 1561 GCGCCGAAAGAGAGCTGGGCGCACGCGCATATAAACCCTCCGAGGGGTCTATGAGTCC 1620
DB 704 GCGCCGAAAGAGAGCTGGGCGCACGCGCATATAAACCCTCCGAGGGGTCTATGAGTCC 763
QY 1621 ATCAGTCTTCTGAGCCCTCTTGGAGCTGAGAGACACCTCAGAGACCAAGTACATGCA 1680
DB 764 ATCAGTCTTCTGAGCCCTCTTGGAGCTGAGAGACACCTCAGAGACCAAGTACATGCA 823
QY 1681 GAGGTCTTCTCTCCACAGAGTCTTACATAGAGAGCCCTCTGACGCCCTTCCACAGCTT 1740

DB 824 GAGGTCTTCTCTCCACAGAGTCTTACATAGAGAGCCCTCTGACGCCCTTCCACAGCTT 883
QY 1741 GCAGGACCCAGAGGAGGCCACAGAGTGTGACAGGGGGGCTCAGCCCTGAGAGTCCCGC 1800
DB 884 GCAGGACCCAGAGGAGGCCACAGAGTGTGACAGGGGGGCTCAGCCCTGAGAGTCCCGC 943
QY 1801 CAGTCAGTGTGTACCTCTCCAGGGCAGTGTGCTGCTGCGCAACCCGACCCAGGCTTCCAG 1860
DB 944 CAGTCAGTGTGTACCTCTCCAGGGCAGTGTGCTGCTGCGCAACCCGACCCAGGCTTCCAG 1003
QY 1861 GAGCAGGACGAGGGGCGAGGGGCGAGGGGCGAGGAGAGCCCTGCAATTTCTCTACGCCCAGG 1920
DB 1004 GAGCAGGACGAGGGGCGAGGGGCGAGGGGCGAGGAGAGCCCTGCAATTTCTCTACGCCCAGG 1063
QY 1921 TTCCGGAAGGTGTGAGACAGGCCAGCGTGATGACAGTGTGAGAGAGGGCGAGGCTGA 1980
DB 1064 TTCCGGAAGGTGTGAGACAGGCCAGCGTGATGACAGTGTGAGAGAGGGCGAGGCTGA 1123
RESULT 12
ABL40801
ID ABL40801 standard; DNA; 1755 BP.
XX
AC ABL40801;
XX
DT 03-JUL-2002 (first entry)
XX
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX
KW Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KW antidiabetic; anorectic; cytostatic; cardiant; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KW neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KW antiallergic; dermatological; vulnerary; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200220732-A2
XX
PD 14-MAR-2002.
XX
PF 27-AUG-2001; 2001WO-EP09848.
XX
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB) BAYER AG.
XX
PI Liou J;
XX
DR MPI; 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
XX
PS Disclosure; Fig 4; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases

CC including myocardial infarction, ischaemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.

Sequence 1755 BP; 350 A; 577 C; 505 G; 323 T; 0 other;

Query Match	35.1%;	Score 694;	DB 24;	Length 1755;
Best Local Similarity	89.1%;	Pred. No. 2.2e-137;		
Matches 802; Conservative	0;	Mismatches 0;	Indels 98;	Gaps 1;

QY	1081	ATGCCCCGGAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCACAAATGTGCCGCTC	1140
Db	322	ATGCCCCGGAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCACAAATGTGCCGCTC	381
QY	1141	TGGGATGAGAGTCGGCCCCAGCTGTGCCCCGCACTGGAAGAGACGCACCGCTTCATTGAG	1200
Db	382	TGGGATGAGAGTCGGCCCCAGCTGTGCCCCGCACTGGAAGAGACGCACCGCTTCATTGAG	441
QY	1201	GCTGCAAGAGCACAGGGGCACCCACTGTGCTGTCCACTGCAAGATGGGCGTCAGCCGCTCA	1260
Db	442	GCTGCAAGAGCACAGGGGCACCCACTGTGCTGTCCACTGCAAGATGGGCGTCAGCCGCTCA	501
QY	1261	GCGGCCACAGTGTGCTGGCTATGCCATGAGAGCACTAACAATGCAGCCTGAGAGAGGCCCTG	1320
Db	502	GCGGCCACAGTGTGCTGGCTATGCCATGAGAGCACTAACAATGCAGCCTGAGAGAGGCCCTG	561
QY	1321	CGCCACGTGACAGAGAGCTCCGGCCCCATCGCCCCGCCCAACCTTGCTTCCTGCGCCAGCTG	1380
Db	562	CGCCACGTGACAGAGAGCTCCGGCCCCATCGCCCCGCCCAACCTTGCTTCCTGCGCCAGCTG	621
QY	1381	CAGATCTACCAGGGGCATCTGCAGCGGCCAGCCGCCAGAGGCCATGTCTGGAGACAGAAAGTG	1440
Db	622	CAGATCTACCAGGGGCATCTGCAGCGGCCAGCCGCCAGAGGCCATGTCTGGAGACAGAAAGTG	645
QY	1441	GGTGGGTCTCCCCAGAGAGACACCAGCCCCCTGAAGTCTCTACACCATTCCCACTCTT	1500
Db	646	-----	645
QY	1501	CCGCAGAACCTGAGGGTGTGGGAGAGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCA	1560
Db	646	--GCCAGAACCTGAGGGTGTGGGAGAGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCA	703
QY	1561	GCCCCGAAAGAGAGAGCTTGGGCCACCGGCCACGTATAAAGCTCCGAGGGGTCAATGAGTCC	1620
Db	704	GCCCCGAAAGAGAGAGCTTGGGCCACCGGCCACGTATAAAGCTCCGAGGGGTCAATGAGTCC	763
QY	1621	ATCAGTCTTCTGAGAGCCCTCTTGAGCTTGAGAGCACCTCAGAGACCAGTGACATGCCA	1680
Db	764	ATCAGTCTTCTGAGAGCCCTCTTGAGCTTGAGAGCACCTCAGAGACCAGTGACATGCCA	823
QY	1681	GAGTCTTCTCTTCCCAAGAGTCTTACATGAAGAGCCTTGACGCCCTTCCCAAGCTT	1740
Db	824	GAGTCTTCTCTTCCCAAGAGTCTTACATGAAGAGCCTTGACGCCCTTCCCAAGCTT	883
QY	1741	GCAAGGACCAAGGGAGGCCAGCAGGTGACAGGGGGCTCAGCCTGCCCCGTAAGTCCGC	1800
Db	884	GCAAGGACCAAGGGAGGCCAGCAGGTGACAGGGGGCTCAGCCTGCCCCGTAAGTCCGC	943
QY	1801	CAGTCAGTGTACCTCTCAGGGCAGTGCCTGTGTGGCCAACCGAACCCAGGCTTCCAG	1860
Db	944	CAGTCAGTGTACCTCTCAGGGCAGTGCCTGTGTGGCCAACCGAACCCAGGCTTCCAG	1003
QY	1861	GAGCAGGAGCAGGGGCGAGGGGCGAGGGGCGAGGAGAGCCCTGCATTCTCTCTACGCCAGG	1920
Db	1004	GAGCAGGAGCAGGGGCGAGGGGCGAGGGGCGAGGAGAGCCCTGCATTCTCTCTACGCCAGG	1063
QY	1921	TTCCGGAAGGTGTGTGAGACAGGCCAGCGTGCATGACAGTGGAGAGAGGGCGAGGCTTGA	1980
Db	1064	TTCCGGAAGGTGTGTGAGACAGGCCAGCGTGCATGACAGTGGAGAGAGGGCGAGGCTTGA	1123

RESULT 13

ID	ABL40803	standard; DNA; 599 BP.
ABL40803		

AC ABL40803;

DT 03-JUL-2002 (first entry)

Human MAP kinase phosphatase-like enzyme DNA fragment.

Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
 antiasthmatic; antidiabetic; anorectic; cyostatic; cardiant; human;
 antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene
 neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
 antiallergic; dermatological; vulnery; gene therapy; ds.

OS Homo sapiens.

PN WO200220732-A2.

PD 14-MAR-2002.

PF 27-AUG-2001; 2001WO-EP09848.

PR 07-SEP-2000; 2000US-230709P.

PA (FARB) BAYER AG.

Liou J; PI

DR WPI; 2002-339802/37.

PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases -

PS Disclosure; Fig 8; 134pp; English.

The invention relates to a purified human mitogen activated protein (MAP) kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed by standard recombinant methodology. The MAP kinase phosphatase-like enzyme and encoding polynucleotides are useful for screening for modulators which are used for treating a MAP kinase phosphatase-like enzyme dysfunction related disease, such as asthma, a central nervous system disorder, diabetes, obesity, chronic obstructive pulmonary disease, cancer or a cardiovascular disease. The enzyme can be regulated to treat allergies including asthma, allergic rhinitis, atopic dermatitis, and anaphylaxis, central nervous system disorders such as brain injuries, Parkinson's disease, dementia, multiple sclerosis, stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human immunodeficiency virus (HIV) dementia, and cardiovascular diseases including myocardial infarction, ischaemic diseases of the heart, atrial and ventricular arrhythmia, hypertensive vascular diseases and peripheral vascular diseases. The enzyme is useful in diagnostic assays for detecting diseases and abnormalities or susceptibility to diseases or abnormalities related to the presence of mutations in the encoding nucleic acid sequences. The present sequence represents the human MAP kinase phosphatase-like enzyme DNA fragment.

Sequence 599 BP; 135 A; 185 C; 183 G; 96 T; 0 other;

Query Match	24.8%;	Score 491;	DB 24;	Length 599;
Best Local Similarity	85.9%;	Pred. No. 1.6e-94;		
Matches 599;	Conservative	0;	Mismatches	0;
			Indels	98;
			Gaps	1;

QY 1024 TGAACGACGCAAACTTGGAGAGCTGCAGAGGAACAGGGTCAACCCATCTTGAACATG 1083

Db 1 TGAACGACGCAAACTTGGAGAGCTGCAGAGGAACAGGGTCAACCCATCTTGAACATG 60

QY 1084 GCGGGGAGATTGACAACTTCTACCTGAGCGCTTCACCTACCAATGTGCGCTCTGG 1143
DB 61 GCGGGGAGATTGACAACTTCTACCTGAGCGCTTCACCTACCAATGTGCGCTCTGG 120
QY 1144 GATGAGAGTTCGCGCCAGCTGTGCTGCCGACTGGAAGAGAGACCGCTTCATTGAGGCT 1203
DB 121 GATGAGAGTTCGCGCCAGCTGTGCTGCCGACTGGAAGAGAGACCGCTTCATTGAGGCT 180
QY 1204 GCAAGAGCACAGGCAACCCACGCTGTGCTGCTCACTGCAAGATGGCGCTCAGCCGCTCAGCG 1263
DB 181 GCAAGAGCACAGGCAACCCACGCTGTGCTGCTCACTGCAAGATGGCGCTCAGCCGCTCAGCG 240
QY 1264 GCCACAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACAGCCTGGAGCAGCGCCCTGCGC 1323
DB 241 GCCACAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACAGCCTGGAGCAGCGCCCTGCGC 300
QY 1324 CACGTGACAGAGCTCCGCGCCCATGCGCGCCCAACCCCTGGCTTCTGCGCCAGCTGACG 1383
DB 301 CACGTGACAGAGCTCCGCGCCCATGCGCGCCCAACCCCTGGCTTCTGCGCCAGCTGACG 360
QY 1384 ATCTACACGAGGATCTGACGCGCCAGCGCCAGACCATGTCTGGAGCAGAAAGTGGGT 1443
DB 361 ATCTACACGAGGATCTGACGCGCGCCAGCGCCAGACCATGTCTGGAGCAGAAAGTGGGT 381
QY 1444 GGGGTCTCCCCAGAGAGACCCAGCCCTGAAGTCTTACACCATTCGCCACTCTCCG 1503
DB 382 -----G 382
QY 1504 CCAGAACCTGAGGGTGTGGGGAGAGAGAGTGTAGGCATGGAAGAGAGCCAGGACGCC 1563
DB 383 CCAGAACCTGAGGGTGTGGGGAGAGAGAGTGTAGGCATGGAAGAGAGCCAGGACGCC 442
QY 1564 CCGAAGAAGAGAGCTGGGGCCACGCGCCACGTAATAACCTCCGAGGGGTCTATGAGGTCAATC 1623
DB 443 CCGAAGAAGAGAGCTGGGGCCACGCGCCACGTAATAACCTCCGAGGGGTCTATGAGGTCAATC 502
QY 1624 AGCTTCTTGAGAGCTCTCTCTTGAGAGCTGAGAGCACCCTCAGAGACCATGACATGCCAGAG 1683
DB 503 AGCTTCTTGAGAGCTCTCTCTTGAGAGCTGAGAGCACCCTCAGAGACCATGACATGCCAGAG 562
QY 1684 GTCTTCTCTTCCACGAGTCTTCACATGAAGCCCTC 1720
DB 563 GTCTTCTCTTCCACGAGTCTTCACATGAAGCCCTC 599
RESULT 14
ABL40802
ID ABL40802 standard; DNA; 409 BP.
AC ABL40802;
XX
DT 03-JUL-2002 (first entry)
DE Human MAP kinase phosphatase-like enzyme DNA fragment.
XX Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;
KW antiasthmatic; antidiabetic; anorectic; cytostatic; cardiant; human;
KW antiparkinsonian; cerebroprotective; neuroprotective; nootropic; gene;
KW neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;
KW antiallergic; dermatological; vulnerary; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200220732-A2.
XX
PD 14-MAR-2002.
XX
PF 27-AUG-2001; 2001WO-EP09848.
XX
PR 07-SEP-2000; 2000US-230709P.
XX
PA (FARB) BAYER AG.
XX

PI Liou J;
XX
DR WPI; 2002-339802/37.
XX
PT New human mitogen activated protein kinase phosphatase-like enzyme
PT polypeptide, regulators of which are useful for preventing, treating
PT allergies including asthma, diabetes, obesity, cancer and
PT cardiovascular diseases
XX
PS Disclosure; Fig 7; 134pp; English.
XX
CC The invention relates to a purified human mitogen activated protein (MAP)
CC kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed
CC by standard recombinant methodology. The MAP kinase phosphatase-like
CC enzyme and encoding polynucleotides are useful for screening for
CC modulators which are used for treating a MAP kinase phosphatase-like
CC enzyme dysfunction related disease, such as asthma, a central nervous
CC system disorder, diabetes, obesity, chronic obstructive pulmonary
CC disease, cancer or a cardiovascular disease. The enzyme can be regulated
CC to treat allergies including asthma, allergic rhinitis, atopic
CC dermatitis, and anaphylaxis, central nervous system disorders such as
CC brain injuries, Parkinson's disease, dementia, multiple sclerosis,
CC stroke, Alzheimer's disease, Huntington's disease, schizophrenia, Pick's
CC disease, Creutzfeldt-Jacob dementia, progressive nuclear palsy, and human
CC immunodeficiency virus (HIV) dementia, and cardiovascular diseases
CC including myocardial infarction, ischaemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases and peripheral
CC vascular diseases. The enzyme is useful in diagnostic assays for
CC detecting diseases and abnormalities or susceptibility to diseases or
CC abnormalities related to the presence of mutations in the encoding
CC nucleic acid sequences. The present sequence represents the human MAP
CC kinase phosphatase-like enzyme DNA fragment.
XX
SQ Sequence 409 BP; 87 A; 127 C; 121 G; 73 T; 1 other;
Query Match 20.1%; Score 397; DB 24; Length 409;
Best Local Similarity 99.5%; Pred. No. 1.2e-74;
Matches 408; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 819 GCGATCCGTGCTGAGCTGTGAAAGTGTGATGTCACTGACCTGGAGAGTCACTTC 878
DB 1 GCGATCCGTGCT-AGCTGTGAAAGTGTGATGTCACTGACCTGGAGAGTCACTTC 59
QY 879 CAAAGATCCGCGCAGCTGTGAGCTGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGA 938
DB 60 CAAAGATCCGCGCAGCTGTGAGCTGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGA 119
QY 939 CTTATCGACAACCAAGATGCTGCTGTGTGACACGCGGAGAGCTCCGCATCTT 998
DB 120 CTTATCGACAACCAAGATGCTGCTGTGTGACACGCGGAGAGCTCCGCATCTT 179
QY 999 CCCCCAACCCTTACCTGGGCTCAGAGTGAACCGCAAAACCTGGAGAGCTGCAGAGAA 1058
DB 180 CCCCCAACCCTTACCTGGGCTCAGAGTGAACCGCAAAACCTGGAGAGCTGCAGAGAA 239
QY 1059 CAGGGTCAACCAATCTTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTT 1118
DB 240 CAGGGTCAACCAATCTTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTT 299
QY 1119 CACCTACCACAATGTGCGCTCTGGGATGAGAGTCCGCCAGCTGTGCCGCACTGGA 1178
DB 300 CACCTACCACAATGTGCGCTCTGGGATGAGAGTCCGCCAGCTGTGCCGCACTGGA 359
QY 1179 GGAGAGCAGCCGCTTCATTGAGGCTGCAAGAGCAGGCGACCCACGCTGC 1228
DB 360 GGAGAGCAGCCGCTTCATTGAGGCTGCAAGAGCAGGCGACCCACGCTGC 409
RESULT 15
AAH98183/C
ID AAH98183 standard; cDNA; 717 BP.
XX
AC AAH98183;

XX 12-OCT-2001 (first entry)
DT Human EST-derived coding sequence SEQ ID NO: 40.
XX
DE Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KM diagnostics; forensic test; gene mapping; genetic disorder;
KM biodiversity; gene therapy; nutrition; ss.
XX
OS Homo sapiens.
XX WO200154477-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US02687.
XX
XX 25-JAN-2000; 2000US-0491404.
XX 17-JUL-2000; 2000US-0617746.
XX 03-AUG-2000; 2000US-0631451.
XX 15-SEP-2000; 2000US-0663870.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
XX WPI; 2001-476164/51.
XX P-PSDB; AAM23524.
XX
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
XX antibodies and research use -
XX
XX Claim 1; Page 210; 1275pp; English.
XX
XX The present invention provides the protein and coding sequences of novel
XX proteins from a variety of organisms, including human, dog, cat, horse,
XX cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX from the organism of interest. They can be used in diagnostics,
XX forensics, gene mapping, identification of mutations, to assess
XX biodiversity and for nutritional purposes. The present sequence is a cDNA
XX of the invention.
XX
SQ Sequence 717 BP; 143 A; 222 C; 200 G; 152 T; 0 other;

Query Match 19.1%; Score 377.2; DB 22; Length 717;
Best Local Similarity 99.2%; Pred. No. 2e-70;
Matches 379; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 87 GAGTCGACTCCAGCAGGAGGAGGAGCTTTGGCGTCTCCGCTGGGCTGTCTGGACTGCA 146
|
DB 697 GGGTCGACTCCAGCAGGAGGAGGAGCTTTGGCGTCTCCGCTGGGCTGTCTGGACTGCA 638
|
QY 147 GGATGAGGGGACATGATGATGACAGAGGAGGAGCTTGTAGCCACAGAGAGAGGCCCC 206
|
DB 637 GGATGAGGGGACATGATGATGACAGAGGAGGAGCTTGTAGCCACAGAGAGAGGCCCC 578
|
QY 207 GAGTGAGAGGAGCTCCACGGGGACACAGACACTTCCGGCAAGAGATCCCA 266
|
DB 577 GAGTGAGAGGAGCTCCACGGGGACACAGACACTTCCGGCAAGAGATCCCA 518
|
QY 267 GAAGCAGAGAGAGCAGAGGAGCAGCCTGCACTCTGTAAGCTGTGAGGCGGAGGA 326
|
DB 517 GAAGCAGAGAGAGCAGAGGAGCAGCCTGCACTCTGTAAGCTGTGAGGCGGAGGA 458
|
QY 327 TGACATCCGCTGCGAGCCAGCTGAGGAGCAGCCCGGCTCCCGGCTACCTGCT 386
|
DB 457 TGACATCCGCTGCGAGCCAGCTGAGGAGCAGCCCGGCTCCCGGCTACCTGCT 398
|
QY 387 GGTAGTTTCTACACGAGAGAGAGGAGGCTGTGAGCCAGATGAGACGGTCTCTGGGCGT 446
|

Db 397 GGTAGTTTCTACACGAGAGAGAGGAGGCTGTGAGCCAGATGAGACGGTCTCTGGGCGT 338
QY 447 GGATTTCCCTGACAGCAGCTCC 468
|
Db 337 GGATTTCCCTGACAGCAGGTTTC 316
|

Search completed: January 15, 2004, 15:28:15
Job time : 545 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 13:13:53 ; Search time 7060 Seconds
(without alignments)
11473.238 Million cell updates/sec

Title: US-09-955-732-1

Perfect score: 1980

Sequence: 1 atggccctgtcaccagtga.....gagagggagggcgagcctga 1980

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1980	100.0	1980	6	AX451343	AX451343 Sequence
2	1978.4	99.9	2781	6	AX099939	AX099939 Sequence
3	1978.4	99.9	2808	9	AK074432	AK074432 Homo sapi
4	1978.4	99.9	2852	6	AX574786	AX574786 Sequence
5	1804.8	91.2	2905	9	AK094226	AK094226 Homo sapi
6	1774	89.6	2704	6	AX574785	AX574785 Sequence
7	1774	89.6	2704	6	AX574791	AX574791 Sequence
8	1772.4	89.5	2604	6	AB072360	AB072360 Homo sapi
9	1684	85.1	2322	6	AX398774	AX398774 Sequence
10	1409	71.2	1416	6	AX451362	AX451362 Sequence
11	1345.6	65.3	2540	6	AX574787	AX574787 Sequence
12	1292.6	65.3	2736	10	BC028922	BC028922 Mus muscu
13	1153.4	58.3	1905	9	BC004176	BC004176 Homo sapi
14	1153.4	58.3	1905	9	BC004210	BC004210 Homo sapi
15	1026	51.8	1026	6	AX086034	AX086034 Sequence
16	851	43.0	2061	6	AX405828	AX405828 Sequence
17	851	43.0	2064	9	AK000522	AK000522 Homo sapi
18	694	35.1	1755	6	AX398768	AX398768 Sequence
19	694	35.1	1755	6	BD156714	BD156714 Primer fo
20	694	35.1	1755	9	AK001790	AK001790 Homo sapi
21	491	24.8	599	6	AX398772	AX398772 Sequence
22	397	20.1	409	6	AX398771	AX398771 Sequence
23	372	18.8	173052	9	AP001885	AP001885 Homo sapi
24	362	18.3	2280	9	AB072356	AB072356 Homo sapi
25	362	18.3	2433	9	AK095421	AK095421 Homo sapi
26	362	18.3	3817	9	AB072355	AB072355 Homo sapi
27	362	18.3	6374	6	AX180874	AX180874 Sequence
28	356.4	18.0	494	6	AX398773	AX398773 Sequence
29	340.4	17.2	1949	6	AX223960	AX223960 Sequence
30	329	16.6	426	6	AX398765	AX398765 Sequence
31	319.2	16.1	3209	10	BC046529	BC046529 Mus muscu
32	308.8	15.6	1711	6	AX223964	AX223964 Sequence
33	306.8	15.5	1796	9	AB072358	AB072358 Homo sapi
34	300.8	15.2	3488	6	AX406972	AX406972 Sequence
35	293.6	14.8	4417	3	AB036834	AB036834 Drosophila
36	282.8	14.3	2260	6	AX180876	AX180876 Sequence
37	279.4	14.1	1771	6	AX202239	AX202239 Sequence
38	256.8	13.0	258710	2	AC135823	AC135823 Rattus no
39	233.4	11.8	571	6	BD149049	BD149049 Primer fo
40	231.8	11.7	302961	2	AC140073	AC140073 Mus muscu
41	204	10.3	1052	6	AX223966	AX223966 Sequence
42	193	9.7	92967	2	AC139555	AC139555 Homo sapi
43	193	9.7	92967	2	AC139555	AC139555 Homo sapi
44	142.8	7.2	386	9	AF484838	AF484838 Homo sapi
45	142	7.2	113474	2	AC013932	AC013932 Drosophila

ALIGNMENTS

RESULT 1
AX451343
LOCUS AX451343 1980 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent WO0224740.
ACCESSION AX451343
VERSION AX451343.1 GI:21698394
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Luche, R.M. and Wei, B.
Dsp-15 dual-specificity phosphatase
Patent: WO 0224740-A 1 28-MAR-2002;
JOURNAL

FEATURES Ceplyr, Inc. (US)
Location/Qualifiers
source 1.1980
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 409 a 616 c 640 g 315 t
ORIGIN

Query Match 100.0%; Score 1980; DB 6; Length 1980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1980; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTGTGTCACAGTACCGCTTGGCCCCCGGCGAGCGGCGCTTCCACGCCCCGTGGG 60
DB 1 ATGGCCCTGTGTCACAGTACCGCTTGGCCCCCGGCGAGCGGCGCTTCCACGCCCCGTGGG 60
QY 61 CCCTGGGACCAAGCGGCTCCAGGAGAGTGAATCCAGCGAAGGCAAGCTTTGGCGTG 120
DB 61 CCCTGGGACCAAGCGGCTCCAGGAGAGTGAATCCAGCGAAGGCAAGCTTTGGCGTG 120
QY 121 CTCCTGGGGCTGTCTCTGGGACTGACAGATGAGGGGCAATGATGATGACAGAGGCC 180
DB 121 CTCCTGGGGCTGTCTCTGGGACTGACAGATGAGGGGCAATGATGATGACAGAGGCC 180
QY 181 AGTTCTGAGCCACAGAGAAGGCCCCGAGTGAAGAGAGAGTCCACGGGACCAAGACAGAC 240
DB 181 AGTTCTGAGCCACAGAGAAGGCCCCGAGTGAAGAGAGAGTCCACGGGACCAAGACAGAC 240
QY 241 TTGGGGCAAGATCCAGAGTCCCGCAGAGCAGAGAGAGAGCAGCAGCAGCAGCAGCCTC 300
DB 241 TTGGGGCAAGATCCAGAGTCCCGCAGAGCAGAGAGAGAGCAGCAGCAGCAGCAGCCTC 300
QY 301 ATGTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 360
DB 301 ATGTACAGCTGTGAGGCGCAGAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 360
QY 361 CGGCTCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAAGAGAGAGTCTGAGC 420
DB 361 CGGCTCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAAGAGAGAGTCTGAGC 420
QY 421 CAGATGAGAGCGGTCTCTGCGGCGTGAATTTCCCTGACAGCAGCTCCCCAGCTGCACC 480
DB 421 CAGATGAGAGCGGTCTCTGCGGCGTGAATTTCCCTGACAGCAGCTCCCCAGCTGCACC 480
QY 481 CTGGGCTGTCTTGGCCCCCTCTGAGTGAACACCCAGGTACTTAGATGAGACGGGGC 540
DB 481 CTGGGCTGTCTTGGCCCCCTCTGAGTGAACACCCAGGTACTTAGATGAGACGGGGC 540
QY 541 TTCAAGCTGACGTCTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
DB 541 TTCAAGCTGACGTCTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
QY 601 TGGGCCAAGCTCCAGGATTTGCAACCAAGCATGTGAGGCAAGCTTAGGCAAGGCTTTGA 660
DB 601 TGGGCCAAGCTCCAGGATTTGCAACCAAGCATGTGAGGCAAGCTTAGGCAAGGCTTTGA 660
QY 661 CCGGCTGAGTGCCTCACTGGGCGCAGCCACTTACAGAGAGAGACTGAATCCGAACAG 720
DB 661 CCGGCTGAGTGCCTCACTGGGCGCAGCCACTTACAGAGAGAGACTGAATCCGAACAG 720
QY 721 AGCTGCTCAATGATGAGCGGCTATGGCCGAGCTGAGTCTTGGCGCTCCAGCGCC 780
DB 721 AGCTGCTCAATGATGAGCGGCTATGGCCGAGCTGAGTCTTGGCGCTCCAGCGCC 780
QY 781 GAGCTGGCGGGTCTCTAGAACAGAGCAGATGAGAGCGCATCCGTGCTGAGCTGTG 840
DB 781 GAGCTGGCGGGTCTCTAGAACAGAGCAGATGAGAGCGCATCCGTGCTGAGCTGTG 840
QY 841 AAAGTGTGATGTAGTGAAGCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGCTCTG 900
DB 841 AAAGTGTGATGTAGTGAAGCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGCTCTG 900

QY 901 GAGCTGGCGCTGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAAACAGATGCTG 960
DB 901 GAGCTGGCGCTGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAAACAGATGCTG 960
QY 961 CTGCTGTGGCACAGCGGACCGAGCCTCCCGCATCTTCCCCCACTTACCTGGGCTCA 1020
DB 961 CTGCTGTGGCACAGCGGACCGAGCCTCCCGCATCTTCCCCCACTTACCTGGGCTCA 1020
QY 1021 GAGTGAACGACGAAACCTGAGAGAGCTGACAGAGAAACAGGGTCAACCATCTTGAAC 1080
DB 1021 GAGTGAACGACGAAACCTGAGAGAGCTGACAGAGAAACAGGGTCAACCATCTTGAAC 1080
QY 1081 ATGGCCCGGAGATTTGACAATTTCTACCTGAGCGCTTCACTTACCAATATGTGCGCTC 1140
DB 1081 ATGGCCCGGAGATTTGACAATTTCTACCTGAGCGCTTCACTTACCAATATGTGCGCTC 1140
QY 1141 TGGGATGAGAGTCCGCCCAAGCTGTGCGGCACTGGAAGGAGAGCAGCCCTTCAATTGAG 1200
DB 1141 TGGGATGAGAGTCCGCCCAAGCTGTGCGGCACTGGAAGGAGAGCAGCCCTTCAATTGAG 1200
QY 1201 GCTGCAAGAGACACAGGGCAACCCAGCTGTGTCTCACTGCAAGATGGCGCTCAGCGCTCA 1260
DB 1201 GCTGCAAGAGACACAGGGCAACCCAGCTGTGTCTCACTGCAAGATGGCGCTCAGCGCTCA 1260
QY 1261 GCGGCCACAGTGTCTGCTTATGCCATGAAAGCAATACGAATGAGCCTGAGCAGCGCTG 1320
DB 1261 GCGGCCACAGTGTCTGCTTATGCCATGAAAGCAATACGAATGAGCCTGAGCAGCGCTG 1320
QY 1321 CGCCACGTGACAGAGCTCCGGCCCATGCCCCCAACCTGTGCTTCTGCGGCAAGCTG 1380
DB 1321 CGCCACGTGACAGAGCTCCGGCCCATGCCCCCAACCTGTGCTTCTGCGGCAAGCTG 1380
QY 1381 CAGATCTACAGGCGATCTGACGCGCACCGCCAGAGCCATGTCTGGAGCAGAAAGTG 1440
DB 1381 CAGATCTACAGGCGATCTGACGCGCACCGCCAGAGCCATGTCTGGAGCAGAAAGTG 1440
QY 1441 GGTGGGTCTCCCAAGAGAGACACCCCTGAAGTCTTCAACCATTCACCTCTT 1500
DB 1441 GGTGGGTCTCCCAAGAGAGACACCCCTGAAGTCTTCAACCATTCACCTCTT 1500
QY 1501 CCGCAGAACTGAGGGTGTGGGAGAGAAAGTTGTAGGCATGGAAGAGCCAGGCA 1560
DB 1501 CCGCAGAACTGAGGGTGTGGGAGAGAAAGTTGTAGGCATGGAAGAGCCAGGCA 1560
QY 1561 GCGCCGAAAGAGAGCTGGGCCAGCGCCACGTATAAACCCTCCGAGGGTCAAGGTCC 1620
DB 1561 GCGCCGAAAGAGAGCTGGGCCAGCGCCACGTATAAACCCTCCGAGGGTCAAGGTCC 1620
QY 1621 ATCAGTCTTGTGAAGCCCTCTGTGAGCTGAGAGCACTCAGAGACCAATGATGCCA 1680
DB 1621 ATCAGTCTTGTGAAGCCCTCTGTGAGCTGAGAGCACTCAGAGACCAATGATGCCA 1680
QY 1681 GAGTCTTCTTCCACAGAGTCTTCAATGAAGAGCTCTGAGCCCTTCCACAGCTT 1740
DB 1681 GAGTCTTCTTCCACAGAGTCTTCAATGAAGAGCTCTGAGCCCTTCCACAGCTT 1740
QY 1741 GCAAGGACCAAGGAGGCGCAGAGGTGACAGGGGCTCAGCTGCTGAAAGTCCCGC 1800
DB 1741 GCAAGGACCAAGGAGGCGCAGAGGTGACAGGGGCTCAGCTGCTGAAAGTCCCGC 1800
QY 1801 CAGTCAAGTGTACCTTCCAGGGCAGTGCCTGTGGCCAAACCGGACCGAGGCTTCCAG 1860
DB 1801 CAGTCAAGTGTACCTTCCAGGGCAGTGCCTGTGGCCAAACCGGACCGAGGCTTCCAG 1860
QY 1861 GAGCAGAGCAGGGGCGAGGGGCGAGGGGAGAGCCCTGCAATTCCTTACGCCCAGG 1920
DB 1861 GAGCAGAGCAGGGGCGAGGGGCGAGGGGAGAGCCCTGCAATTCCTTACGCCCAGG 1920
QY 1921 TTCCGGAAGTGTGAGACAGGCCAGCGTGAATGACAGTGAAGAGAGGCGAGGCTTGA 1980
DB 1921 TTCCGGAAGTGTGAGACAGGCCAGCGTGAATGACAGTGAAGAGAGGCGAGGCTTGA 1980

```
RESULT 2
AX099939 2781 bp DNA linear PAT 02-APR-2001
LOCUS Sequence 21 from Patent WO0120004.
DEFINITION AX099939
ACCESSION AX099939
VERSION AX099939.1 GI:13538949
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Yue,H., Tang,Y.T., Bandman,O., Hillman,J.L., Baughn,M.R.,
Azimzai,Y. and Lu,D.A.
Protein phosphatase and kinase proteins
Patent: WO 0120004-A 21 22-MAR-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 2781
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID NO: 5039718CB1"
BASE COUNT 576 a 906 c 820 g 479 t
ORIGIN
Query Match 99.9%; Score 1978.4; DB 6; Length 2781;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCCCCGTGTACAGTGAAGCCGTTGCCCCCGGCGAGCGGCGCTTCCACGCCCGTGCGG 60
Db ATGCCCTGTGTACAGTGAAGCCGTTGCCCCCGGCGAGCGGCGCTTCCACGCCCGTGCGG 146
QY 61 CCTGGGACCCAGGCGGTCCAGCGAAGAGTGCATCTCCAGCGAAGGACAGCTTTGGGTG 120
Db CCTGGGACCCAGGCGGTCCAGCGAAGAGTGCATCTCCAGCGAAGGACAGCTTTGGGTG 206
QY 121 CTCCTGGGGGTGTCTCTGGGACTGCAGATGAGGGGACATGATGACAGAGGCC 180
Db CTCCTGGGGGTGTCTCTGGGACTGCAGATGAGGGGACATGATGACAGAGGCC 266
QY 207 CTCCTGGGGGTGTCTCTGGGACTGCAGATGAGGGGACATGATGACAGAGGCC 266
Db CTCCTGGGGGTGTCTCTGGGACTGCAGATGAGGGGACATGATGACAGAGGCC 326
QY 181 AGTTCTGAGCCCAACAGAGAAGGCCCCAGTGAGAGAGAGCTCCACGGGACAGACAGAC 240
Db AGTTCTGAGCCCAACAGAGAAGGCCCCAGTGAGAGAGAGCTCCACGGGACAGACAGAC 326
QY 267 AGTTCTGAGCCCAACAGAGAAGGCCCCAGTGAGAGAGAGCTCCACGGGACAGACAGAC 326
Db AGTTCTGAGCCCAACAGAGAAGGCCCCAGTGAGAGAGAGCTCCACGGGACAGACAGAC 326
QY 241 TTCGGGCAAGGATCCCAAGATCCCCAGAGAGAGAGAGAGAGGACAGCACTGCACCTC 300
Db TTCGGGCAAGGATCCCAAGATCCCCAGAGAGAGAGAGAGAGGACAGCACTGCACCTC 386
QY 327 TTCGGGCAAGGATCCCAAGATCCCCAGAGAGAGAGAGAGAGGACAGCACTGCACCTC 386
Db TTCGGGCAAGGATCCCAAGATCCCCAGAGAGAGAGAGAGAGGACAGCACTGCACCTC 446
QY 301 ATGCTACAGCTGTGAGGCGCGAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 360
Db ATGCTACAGCTGTGAGGCGCGAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 446
QY 387 ATGCTACAGCTGTGAGGCGCGAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 446
Db ATGCTACAGCTGTGAGGCGCGAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 446
QY 361 CGGCTCCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAGAGAGGCTGTGAGC 420
Db CGGCTCCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAGAGAGGCTGTGAGC 506
QY 447 CGGCTCCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAGAGAGGCTGTGAGC 506
Db CGGCTCCCCGGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAGAGAGGCTGTGAGC 566
QY 421 CAGATGAGACGGTCTCTCTGGGGGTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACC 480
Db CAGATGAGACGGTCTCTCTGGGGGTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACC 566
QY 507 CAGATGAGACGGTCTCTCTGGGGGTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACC 566
Db CAGATGAGACGGTCTCTCTGGGGGTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACC 566
QY 481 CTGGGCGCTGTCTTGGCCCCCTCTGAGTGAACACCCAGGTGTAATTAGATGAGACGGGGGC 540
Db CTGGGCGCTGTCTTGGCCCCCTCTGAGTGAACACCCAGGTGTAATTAGATGAGACGGGGGC 626
QY 567 CTGGGCGCTGTCTTGGCCCCCTCTGAGTGAACACCCAGGTGTAATTAGATGAGACGGGGGC 626
Db CTGGGCGCTGTCTTGGCCCCCTCTGAGTGAACACCCAGGTGTAATTAGATGAGACGGGGGC 686
QY 541 TTCAGCGTGAGCTGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
Db TTCAGCGTGAGCTGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 686
QY 601 TGGGCAACACTCCAGGTAATTGACCAAGCATGTGAGGAGAGCTTAGGACGGGCTTGTGTA 660
Db TGGGCAACACTCCAGGTAATTGACCAAGCATGTGAGGAGAGCTTAGGACGGGCTTGTGTA 746
```

```
QY 661 CCGGTGGCAGTGCCTCTCACTGGGCGAGCCCACTACAGAGAGAGTGAATCCGAACAG 720
Db CCGGTGGCAGTGCCTCTCACTGGGCGAGCCCACTACAGAGAGAGTGAATCCGAACAG 806
QY 721 AGTGCCTCAATGATGATGAGCGGCTATGGCCGACCTGAGTCTGTGGGCTCCAGCGCC 780
Db AGTGCCTCAATGATGATGAGCGGCTATGGCCGACCTGAGTCTGTGGGCTCCAGCGCC 866
QY 807 AGTGCCTCAATGATGATGAGCGGCTATGGCCGACCTGAGTCTGTGGGCTCCAGCGCC 866
Db AGTGCCTCAATGATGATGAGCGGCTATGGCCGACCTGAGTCTGTGGGCTCCAGCGCC 866
QY 781 GAGCTGGCGGGTCTCTCAGAACAGAGCAGATGAGACAGCGGATCCGTGTAGCTGTGG 840
Db GAGCTGGCGGGTCTCTCAGAACAGAGCAGATGAGACAGCGGATCCGTGTAGCTGTGG 926
QY 867 GAGCTGGCGGGTCTCTCAGAACAGAGCAGATGAGACAGCGGATCCGTGTAGCTGTGG 926
Db GAGCTGGCGGGTCTCTCAGAACAGAGCAGATGAGACAGCGGATCCGTGTAGCTGTGG 926
QY 841 AAAGTTGGATGTCACTGAGCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 900
Db AAAGTTGGATGTCACTGAGCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 986
QY 927 AAAGTTGGATGTCACTGAGCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 986
Db AAAGTTGGATGTCACTGAGCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 986
QY 901 GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCGGTGACTTCAATGACACAGATGCTG 960
Db GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCGGTGACTTCAATGACACAGATGCTG 1046
QY 987 GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCGGTGACTTCAATGACACAGATGCTG 1046
Db GAGCTGGCGCTGGGGCTCCCTCCAGCAGTACCGGTGACTTCAATGACACAGATGCTG 1046
QY 1047 CTGCTGTGTGACACAGCGGAGCCGAGCTCCCGCATCTTCCCCACCTCTACCTGGGCTCA 1020
Db CTGCTGTGTGACACAGCGGAGCCGAGCTCCCGCATCTTCCCCACCTCTACCTGGGCTCA 1106
QY 1021 GAGTGAACCCAGCAAACTTGAAGAGAGTGCAGAGAGAAAGGGTCAACCCACATCTGAAC 1080
Db GAGTGAACCCAGCAAACTTGAAGAGAGTGCAGAGAGAAAGGGTCAACCCACATCTGAAC 1166
QY 1107 GAGTGAACCCAGCAAACTTGAAGAGAGTGCAGAGAGAAAGGGTCAACCCACATCTGAAC 1166
Db GAGTGAACCCAGCAAACTTGAAGAGAGTGCAGAGAGAAAGGGTCAACCCACATCTGAAC 1166
QY 1081 ATGCCCCGGAGATTGACAACTTTTACCCTGAGCGCTTCACTACCAATGTGCGCTC 1140
Db ATGCCCCGGAGATTGACAACTTTTACCCTGAGCGCTTCACTACCAATGTGCGCTC 1226
QY 1167 ATGCCCCGGAGATTGACAACTTTTACCCTGAGCGCTTCACTACCAATGTGCGCTC 1226
Db ATGCCCCGGAGATTGACAACTTTTACCCTGAGCGCTTCACTACCAATGTGCGCTC 1226
QY 1141 TGGATGAGAGTGGGCCAGCTGTGCGCACTGGAAGAGAGACGACCCGCTTCAATTGAG 1200
Db TGGATGAGAGTGGGCCAGCTGTGCGCACTGGAAGAGAGACGACCCGCTTCAATTGAG 1286
QY 1227 TGGATGAGAGTGGGCCAGCTGTGCGCACTGGAAGAGAGACGACCCGCTTCAATTGAG 1286
Db TGGATGAGAGTGGGCCAGCTGTGCGCACTGGAAGAGAGACGACCCGCTTCAATTGAG 1286
QY 1201 GCTGCAAGAGCAGAGGCAACCCAGCTGTGCTGTCCACTGCAAGATGGGCGTCAAGCCCTCA 1260
Db GCTGCAAGAGCAGAGGCAACCCAGCTGTGCTGTCCACTGCAAGATGGGCGTCAAGCCCTCA 1346
QY 1287 GCTGCAAGAGCAGAGGCAACCCAGCTGTGCTGTCCACTGCAAGATGGGCGTCAAGCCCTCA 1346
Db GCTGCAAGAGCAGAGGCAACCCAGCTGTGCTGTCCACTGCAAGATGGGCGTCAAGCCCTCA 1346
QY 1261 GCGGCCACAGTGTGCTGCTATGCAATGAAGCAGTACGAATGCAAGCTTGAGAGAGCCCTG 1320
Db GCGGCCACAGTGTGCTGCTATGCAATGAAGCAGTACGAATGCAAGCTTGAGAGAGCCCTG 1406
QY 1347 GCGGCCACAGTGTGCTGCTATGCAATGAAGCAGTACGAATGCAAGCTTGAGAGAGCCCTG 1406
Db GCGGCCACAGTGTGCTGCTATGCAATGAAGCAGTACGAATGCAAGCTTGAGAGAGCCCTG 1406
QY 1321 CGCCACGTGAGAGAGCTCCGCGCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1380
Db CGCCACGTGAGAGAGCTCCGCGCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1466
QY 1407 CGCCACGTGAGAGAGCTCCGCGCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1466
Db CGCCACGTGAGAGAGCTCCGCGCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1466
QY 1381 CAGATCTACAGGGCATCTCTGACGGCGACCGCCAGAGCCATGTCTGGAGAGAGAAAGTG 1440
Db CAGATCTACAGGGCATCTCTGACGGCGACCGCCAGAGCCATGTCTGGAGAGAGAAAGTG 1526
QY 1467 CAGATCTACAGGGCATCTCTGACGGCGACCGCCAGAGCCATGTCTGGAGAGAGAAAGTG 1526
Db CAGATCTACAGGGCATCTCTGACGGCGACCGCCAGAGCCATGTCTGGAGAGAGAAAGTG 1526
QY 1441 GGTGGGTCTCCCAAGAGAGAGACCCAGCCCTGAAGTCTCTACACCATTTCCACCTCTT 1500
Db GGTGGGTCTCCCAAGAGAGAGACCCAGCCCTGAAGTCTCTACACCATTTCCACCTCTT 1586
QY 1527 GGTGGGTCTCCCAAGAGAGAGACCCAGCCCTGAAGTCTCTACACCATTTCCACCTCTT 1586
Db GGTGGGTCTCCCAAGAGAGAGACCCAGCCCTGAAGTCTCTACACCATTTCCACCTCTT 1586
QY 1501 CCGCCAGAACCTTGAAGGTGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGAGAGGCA 1560
Db CCGCCAGAACCTTGAAGGTGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGAGAGGCA 1646
QY 1587 CCGCCAGAACCTTGAAGGTGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGAGAGGCA 1646
Db CCGCCAGAACCTTGAAGGTGTGGGAGAGAGAGGTTGTAGGCATGGAAGAGAGAGAGGCA 1646
QY 1561 GCGCCGAAAGAGAGAGCTTGGGCGACGGCCACGCTATAAAGCTCCGAGGGGTATGAGGTCC 1620
Db GCGCCGAAAGAGAGAGCTTGGGCGACGGGCGACGCTATAAAGCTCCGAGGGGTATGAGGTCC 1706
QY 1647 GCGCCGAAAGAGAGAGCTTGGGCGACGGGCGACGCTATAAAGCTCCGAGGGGTATGAGGTCC 1706
Db GCGCCGAAAGAGAGAGCTTGGGCGACGGGCGACGCTATAAAGCTCCGAGGGGTATGAGGTCC 1766
QY 1621 ATCAGTCTTCTGAGAGCCCTCTTGTGAGCTGAGAGACCTTCAAGAGACCAAGTGAATGCCA 1680
Db ATCAGTCTTCTGAGAGCCCTCTTGTGAGCTGAGAGACCTTCAAGAGACCAAGTGAATGCCA 1766
QY 1707 ATCAGTCTTCTGAGAGCCCTCTTGTGAGCTGAGAGACCTTCAAGAGACCAAGTGAATGCCA 1766
Db ATCAGTCTTCTGAGAGCCCTCTTGTGAGCTGAGAGACCTTCAAGAGACCAAGTGAATGCCA 1766
QY 1681 GAGGTCTTCTTCCCAAGAGTCTTCAATGAAGAGCCCTTGACAGCCCTTCCACAGCTT 1740
Db GAGGTCTTCTTCCCAAGAGTCTTCAATGAAGAGCCCTTGACAGCCCTTCCACAGCTT 1826
QY 1767 GAGGTCTTCTTCCCAAGAGTCTTCAATGAAGAGCCCTTGACAGCCCTTCCACAGCTT 1826
Db GAGGTCTTCTTCCCAAGAGTCTTCAATGAAGAGCCCTTGACAGCCCTTCCACAGCTT 1826
```

QY 1741 GCAAGGACCAAGGAGGCCAGCAGGTGACAGGGGGCCCTACGCTGCCCCGTAAGTCCCGC 1800
DB 1827 GCAAGGACCAAGGAGGCCAGCAGGTGACAGGGGGCCCTACGCTGCCCCGTAAGTCCCGC 1886
QY 1801 CAGTCAGTGTACCTCCAGGCGAGTCCGTGTGSCCAACCGGACCGGCTTCCAG 1860
DB 1887 CAGTCAGTGTACCTCCAGGCGAGTCCGTGTGSCCAACCGGACCGGCTTCCAG 1946
QY 1861 GAGCAGGAGCAGGGGCGAGGGGCGAGAGAGCCCTGCAATTCCTCTACGCCCCAGG 1920
DB 1947 GAGCAGGAGCAGGGGCGAGGGGCGAGAGAGCCCTGCAATTCCTCTACGCCCCAGG 2006
QY 1921 TTCGGAAGGTGTGAGACAGGCCGCTGCATGACAGTGAGAGAGGGCGAGGCTTGA 1980
DB 2007 TTCGGAAGGTGTGAGACAGGCCGCTGCATGACAGTGAGAGAGGGCGAGGCTTGA 2066

RESULT 3
AK074432 2808 bp mRNA linear PRI 15-FEB-2002
LOCUS AK074432
DEFINITION Homo sapiens cDNA FLJ23852 fis, clone KAT12021.
ACCESSION AK074432
VERSION AK074432.1 GI:18677033
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2808)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction,
5'- & 3'-end one pass sequencing: Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
source
1..2808
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KAT12021"
/cell_line="KATO III"
/cell_type="signal-ring cell carcinoma"
/clone_lib="KAT"
/note="cloning vector: PME18sFL3"
97..2076
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB85080.1"
/db_xref="GI:18677034"

CDS
/translation="MALIVTSRSPSGASTPVGMQAVORRSLQRQSFALVGA
VIGLQDGDNDAAEASSEPTKAPSEBELHGDQTFGQSGSPQKQEBQRHLHMV
QLLRPQDDIRLAQLEAPRPRLRYLVSTRBEGLSQDETLLGVDPDSSPSPECT
LGLVPLMSDTQVLDGDFSVTSGGQRIKPISTQTMATLQVLHQAEEALGSG
LVPGSALTWASHYGERLNSQSCLENTAMADLESIRPSAEBGSSSEQMEQAIR
AEIMKVLDSVLSVTSKRIQALELRGLPQYRDFIDNQMLLVLAQRDRASRLFP
HLVLSGEMNANLEELQNRVTHILMAREIDNFYPERFTYHNVRLWDESAQLPHM
KETHRFIEARAQGIHVLVHCKMGVSRSAATVLAAMKQYECSEQLRHHVQELPDA
RNPGRFLRLQIYIGILTAQRQSHVMEQKVGVSPEHPAPEVSTPPPLPPEPBG

EEKTVMEESQAPKEPDRPRINLRGMRSISLLESLLESTSETSDMEVFSH
ESSHEBQPPOLARTKGQOVDRGPOPAKRSQSVTLQGSAAVANKTQAFOEQEO
GGGGGGEPCISSTPRSRKVVRAQSVHDSGEERA"
BASE COUNT 585 a 912 c 823 g 488 t
ORIGIN

Query Match 99.9%; Score 1978.4; DB 9; Length 2808;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCCTGGTCAAGTGAAGCCGTTGCCCCCGGGCAGCGGCGCTCCACGCCCCGTGGGG 60
DB 97 ATGGCCCTGGTCAAGTGAAGCCGTTGCCCCCGGGCAGCGGCGCTCCACGCCCCGTGGGG 156
QY 61 CCTGGACCAAGCGGCTCCAGCGAAGAGTGAAGTCCAGCGAAGGAGAGGCTTTGGCGTG 120
DB 157 CCTGGACCAAGCGGCTCCAGCGAAGAGTGAAGTCCAGCGAAGGAGAGGCTTTGGCGTG 216
QY 121 CTCGGTGGGCTGTCTCTGGGACTGACAGATGAGAGGGGACATGATGACAGAGGCC 180
DB 217 CTCGGTGGGCTGTCTCTGGGACTGACAGATGAGAGGGGACATGATGACAGAGGCC 276
QY 181 AGTTCTGAGCCACAGAGAAGGCCCGAGTGAAGAGAGAGGAGGAGGAGGAGGAGGAGG 240
DB 277 AGTTCTGAGCCACAGAGAAGGCCCGAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGG 336
QY 241 TTCGGCAAGATCCACAGAGTCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 337 TTCGGCAAGATCCACAGAGTCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
QY 301 ATGTACAGCTGTGAGAGCGCCGAGAGATGACATCCGCTGGCAGACCCAGCTGAGGACACC 360
DB 397 ATGTACAGCTGTGAGAGCGCCGAGAGATGACATCCGCTGGCAGACCCAGCTGAGGACACC 456
QY 361 CGGCTCCCGCGCTCCGCTACCTGCTGTAGTTTCTACACGAGAAGAGAGAGAGAGAGAG 420
DB 457 CGGCTCCCGCGCTCCGCTACCTGCTGTAGTTTCTACACGAGAAGAGAGAGAGAGAGAG 516
QY 421 CAGGATGAGAGCGTCTCTCTGGGCGTGATTTCCCTGACAGCAGCTCCCGCAGCTGCACC 480
DB 517 CAGGATGAGAGCGTCTCTCTGGGCGTGATTTCCCTGACAGCAGCTCCCGCAGCTGCACC 576
QY 481 CTGGGCTGTGTCTTGGCCCTCTGAGTGAACACCCAGGCTGTAAGATGAGAGCGGGGGC 540
DB 577 CTGGGCTGTGTCTTGGCCCTCTGAGTGAACACCCAGGCTGTAAGATGAGAGCGGGGGC 636
QY 541 TTCAGCGTGAAGTGTGTGGGCAAGCCGAGTCTTAAAGCCATCTCCATCCAGACCATG 600
DB 637 TTCAGCGTGAAGTGTGTGGGCAAGCCGAGTCTTAAAGCCATCTCCATCCAGACCATG 696
QY 601 TGGCCACACATCCAGGTATTGACCAAGCATGTGAGGAGCTCTAGGACGGGCTTGTGTA 660
DB 697 TGGCCACACATCCAGGTATTGACCAAGCATGTGAGGAGCTCTAGGACGGGCTTGTGTA 756
QY 661 CCGGCTGGCAGTGCCTCACCCTGAGCCAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 757 CCGGCTGGCAGTGCCTCACCCTGAGCCAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAG 816
QY 721 AGCTGCTCAATGAGTGAAGCGCTATGCGCAGCTGAGTCTGCGGCTCCAGCGCC 780
DB 817 AGCTGCTCAATGAGTGAAGCGCTATGCGCAGCTGAGTCTGCGGCTCCAGCGCC 876
QY 781 GAGCTGGCGGCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 877 GAGCTGGCGGCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 936
QY 841 AAAGTGTGATGTCAAGTGAAGAGTGTCACTTCAAAGAGATCCGCGCAGGCTCTG 900
DB 937 AAAGTGTGATGTCAAGTGAAGAGTGTCACTTCAAAGAGATCCGCGCAGGCTCTG 996
QY 901 GAGCTGGCGGCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 997 GAGCTGGCGGCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1056

QY	961	CTGCTGTGGCACAGCGGGACCGAGCCTCCCGCATCTTCCCCCACCCTCTACCTGGGCTCA	1020
Db	1057	CTGCTGTGGCACAGCGGGACCGAGCCTCCCGCATCTTCCCCCACCCTCTACCTGGGCTCA	1116
QY	1021	GAGTGAACCGACGAACCTTGAGAGAGCTGCAGAGGAACAGGGTCAACCATCTTGAAC	1080
Db	1117	GAGTGAACCGACGAACCTTGAGAGAGCTGCAGAGGAACAGGGTCAACCATCTTGAAC	1176
QY	1081	ATGCCCCGGGAGATTGACAACTTCTACCTGAGCGCTTCACTTCAACATGTGGCCTC	1140
Db	1177	ATGCCCCGGGAGATTGACAACTTCTACCTGAGCGCTTCACTTCAACATGTGGCCTC	1236
QY	1141	TGGGATGAGGAGTGGCCCCAGCTGCTGCCGACTGGAGAGAGACGCAACGCTTCAATTGAG	1200
Db	1237	TGGGATGAGGAGTGGCCCCAGCTGCTGCCGACTGGAGAGAGACGCAACGCTTCAATTGAG	1296
QY	1201	GCTGCAAGAGCACAGGGGACCCACGCTGCTGTCCACTGCAAGATGGGGCTCAGCCGCTCA	1260
Db	1297	GCTGCAAGAGCACAGGGGACCCACGCTGCTGTCCACTGCAAGATGGGGCTCAGCCGCTCA	1356
QY	1261	GCGGCCACAGTGTGCTGCTATGCTCATGAAAGCAGTACGAATGACGCTGAGCAGGCCCTG	1320
Db	1357	GCGGCCACAGTGTGCTGCTATGCTCATGAAAGCAGTACGAATGACGCTGAGCAGGCCCTG	1416
QY	1321	CGCCACGTGCAGAGCTCCGGCCCATCGCCGCCCAACCCCTGCTTCTGCGCCAGCTG	1380
Db	1417	CGCCACGTGCAGAGCTCCGGCCCATCGCCGCCCAACCCCTGCTTCTGCGCCAGCTG	1476
QY	1381	CAGATCTACAGAGGATCTGACGGCCAGCCGACAGCCATGTCTGGAGCAGAAAGTG	1440
Db	1477	CAGATCTACAGAGGATCTGACGGCCAGCCGACAGCCATGTCTGGAGCAGAAAGTG	1536
QY	1441	GGTGGGGTCTCCCAAGAGACACCCAGCCCTGAAGTCTCTACACCATTCCACCTCTT	1500
Db	1537	GGTGGGGTCTCCCAAGAGACACCCAGCCCTGAAGTCTCTACACCATTCCACCTCTT	1596
QY	1501	CCGCCAGAACCTGAGGGTGTGGGGAGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCA	1560
Db	1597	CCGCCAGAACCTGAGGGTGTGGGGAGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCA	1656
QY	1561	GCCCCGAAAGAGAGGCTGGGCCACGCGCCACGTATAAACCTCCGAGGGGTCTAGAGTCC	1620
Db	1657	GCCCCGAAAGAGAGGCTGGGCCACGCGCCACGTATAAACCTCCGAGGGGTCTAGAGTCC	1716
QY	1621	ATCAGTCTTCTGAGGCTCTTGGAGCTGGAGAGCACCTCAGAGACCAAGTACATGCCA	1680
Db	1717	ATCAGTCTTCTGAGGCTCTTGGAGCTGGAGAGCACCTCAGAGACCAAGTACATGCCA	1776
QY	1681	GAGGCTTCTTCTCCCAAGAGTCTTCAATGAAGAGCCTCTGACGCTTCCACAGCTT	1740
Db	1777	GAGGCTTCTTCTCCCAAGAGTCTTCAATGAAGAGCCTCTGACGCTTCCACAGCTT	1836
QY	1741	GCAAGGACCAAGGGAGGCGCAGAGTGGACAGGGGGCTCAGGCTGCGCTGAAGTCCGC	1800
Db	1837	GCAAGGACCAAGGGAGGCGCAGAGTGGACAGGGGGCTCAGGCTGCGCTGAAGTCCGC	1896
QY	1801	CAGTCAGTGTTACCTTCCAGGGCAGTGCCTGTGGCCCAACCGGACCCAGGCTTCCAG	1860
Db	1897	CAGTCAGTGTTACCTTCCAGGGCAGTGCCTGTGGCCCAACCGGACCCAGGCTTCCAG	1956
QY	1861	GAGCAGAGCAGGGGCGAGGGGCGAGGGGAGAGCCCTGCAATTCTCTACGCCCAAG	1920
Db	1957	GAGCAGAGCAGGGGCGAGGGGCGAGGGGAGAGCCCTGCAATTCTCTACGCCCAAG	2016
QY	1921	TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGCGAGGCTGA	1980
Db	2017	TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGCGAGGCTGA	2076

RESULT 4
AX574786 2852 bp DNA linear PAT 07-JAN-2003
LOCUS AX574786

DEFINITION		Sequence 2 from Patent WO0242436.	
ACCESSION		AX574786	
VERSION		AX574786.1	
KEYWORDS		GI:27551935	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Wei, M.H., Ketchum, K.A., di Francesco, V. and Beasley, E.M. Isolated human phosphatase proteins, nucleic acid molecules encoding human phosphatase proteins, and uses thereof Patent: WO 0242436-A 2 30-MAY-2002; PE Corporation (NY) (US)	
JOURNAL		Location/Qualifiers	
FEATURES		1..2852	
source		/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"	
BASE COUNT		674 a 895 c 807 g 476 t	
ORIGIN			
Query Match		99.9%; Score 1978.4; DB 6; Length 2852;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 1979; Conservative		0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	ATGGCCCTGGTCAACAGTAGCCGTTCCGCCCCGGGAGCGGCGCTCCAGCCCGCTGGGG	60
Db	57	ATGGCCCTGGTCAACAGTAGCCGTTCCGCCCCGGGAGCGGCGCTCCAGCCCGCTGGGG	116
QY	61	CCCTGGACCAAGCGGTCCAGGAGAGAGTGCATCCAGGAAGCAGAGCTTTGGGTG	120
Db	117	CCCTGGACCAAGCGGTCCAGGAGAGAGTGCATCCAGGAAGCAGAGCTTTGGGTG	176
QY	121	CTCCGTGGGGTGTCTCTGGGACTGAGATGAGGGGACAATGATGACAGAGGCC	180
Db	177	CTCCGTGGGGTGTCTCTGGGACTGAGATGAGGGGACAATGATGACAGAGGCC	236
QY	181	AGTTCTAGCCACAGAGAGGCCCGAGTGAAGAGAGCTCCAGCGGACCAAGACAGAC	240
Db	237	AGTTCTAGCCACAGAGAGGCCCGAGTGAAGAGAGCTCCAGCGGACCAAGACAGAC	296
QY	241	TTCCGCAAGATCCCAAGTCCCAAGACAGAGAGAGCAGAGGACAGCCTGCACCTC	300
Db	297	TTCCGCAAGATCCCAAGTCCCAAGACAGAGAGAGCAGAGGACAGCCTGCACCTC	356
QY	301	ATGATAGCTGTGAGGCGCAGAGATGACATCCGCTGGACGCCAGCTGAGGCAACC	360
Db	357	ATGATAGCTGTGAGGCGCAGAGATGACATCCGCTGGACGCCAGCTGAGGCAACC	416
QY	361	CGGCTCCCCGGCTCCGCTACCTGCTGTAGTTCTACAGAGAGAGAGTCTGAGC	420
Db	417	CGGCTCCCCGGCTCCGCTACCTGCTGTAGTTCTACAGAGAGAGAGTCTGAGC	476
QY	421	CAGATGAGACGGTCTCTGGGGCTGGATTTCCTGACAGAGCTCCCCAAGCTGCACC	480
Db	477	CAGATGAGACGGTCTCTGGGGCTGGATTTCCTGACAGAGCTCCCCAAGCTGCACC	536
QY	481	CTGGGCTGTGTTGCCCCCTCTGAGTGACACCCAGGTGTAAGTGAAGAGCGGGGC	540
Db	537	CTGGGCTGTGTTGCCCCCTCTGAGTGACACCCAGGTGTAAGTGAAGAGCGGGGC	596
QY	541	TTACAGGTGACGTCTGTGGGCAAAAGCCGGATCTTCAAGCCCATCTCCAGACCATG	600
Db	597	TTACAGGTGACGTCTGTGGGCAAAAGCCGGATCTTCAAGCCCATCTCCAGACCATG	656
QY	601	TGGGCAACACTCCAGGTAATGCAACCAAGCATGTGAGGAGAGCTTAGGAGCGGCTTGA	660
Db	657	TGGGCAACACTCCAGGTAATGCAACCAAGCATGTGAGGAGAGCTTAGGAGCGGCTTGA	716
QY	661	CCGGTGGCAGTGCCTTCACTGGGCCAGGCACTACAGAGAGAGAGTGAATCCGAACAG	720
Db	717	CCGGTGGCAGTGCCTTCACTGGGCCAGGCACTACAGAGAGAGAGTGAATCCGAACAG	776

QY 721 AGCTGCTCATATGATGAGCGGCTATGGCCGACCTGGAGTCTCTGGGCTCCAGCGCC 780
DB 777 AGCTGCTCATATGATGAGCGGCTATGGCCGACCTGGAGTCTCTGGGCTCCAGCGCC 836
QY 781 GAGCTGGCGGGTCTCTCAGAACGAGCAGATGAGCAGCGGATCCGTGCTGAGCTGTG 840
DB 837 GAGCTGGCGGGTCTCTCAGAACGAGCAGATGAGCAGCGGATCCGTGCTGAGCTGTG 896
QY 841 AAAGTGTGATGTCACTGAGCTGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 900
DB 897 AAAGTGTGATGTCACTGAGCTGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG 956
QY 901 GAGCTGGCGCTGGGCTCCCTCCAGCAGTACCGTGACTTCACTGACCAACAGATGCTG 960
DB 957 GAGCTGGCGCTGGGCTCCCTCCAGCAGTACCGTGACTTCACTGACCAACAGATGCTG 1016
QY 961 CTGCTGTGGACACAGCGGAGCGGAGCTCCCGCATCTTCCCGCCTTACCTGGGCTCA 1020
DB 1017 CTGCTGTGGACACAGCGGAGCGGAGCTCCCGCATCTTCCCGCCTTACCTGGGCTCA 1076
QY 1021 GAGTGAACGACCAACCTGAGAGCTGACAGAGAACAGGTCACCCACATCTTGAAC 1080
DB 1077 GAGTGAACGACCAACCTGAGAGCTGACAGAGAACAGGTCACCCACATCTTGAAC 1136
QY 1081 ATGGCCCGGAGATTTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGCGCTC 1140
DB 1137 ATGGCCCGGAGATTTGACAACTTCTACCTGAGCGCTTACCTACCAATGTGCGCTC 1196
QY 1141 TGGGATGAGAGTCCGGCCAGCTGCTGCCGCACTGGAAGAGACGACCGCTTCAATTGAG 1200
DB 1197 TGGGATGAGAGTCCGGCCAGCTGCTGCCGCACTGGAAGAGACGACCGCTTCAATTGAG 1256
QY 1201 GCTGCAAGAGACAGGGGCAACCCAGCTGCTGTCCACTGCAAGATGGGCGTCAAGCCGCTCA 1260
DB 1257 GCTGCAAGAGACAGGGGCAACCCAGCTGCTGTCCACTGCAAGATGGGCGTCAAGCCGCTCA 1316
QY 1261 GCGGCCACAGTGTGCTGCTATGCTATGAAAGCAAGTACGAATGCAAGCTTGAGCAGGCCCTG 1320
DB 1317 GCGGCCACAGTGTGCTGCTATGCTATGAAAGCAAGTACGAATGCAAGCTTGAGCAGGCCCTG 1376
QY 1321 CGCCACGTGACAGAGCTCCGGGCCATGCGCCGCCAACCTTGCTTCTGCGCCAGCTG 1380
DB 1377 CGCCACGTGACAGAGCTCCGGGCCATGCGCCGCCAACCTTGCTTCTGCGCCAGCTG 1436
QY 1381 CAGATCTACAGAGGCTCTGACCGGCCAGCGCCAGAGCCATGTCTGGAGCAGAAAGTG 1440
DB 1437 CAGATCTACAGAGGCTCTGACCGGCCAGCGCCAGAGCCATGTCTGGAGCAGAAAGTG 1496
QY 1441 GGTGGGTCTCTCCAGAGAGCACCAGCCCTGAAGTCTCTACACCATTTCCACTCTT 1500
DB 1497 GGTGGGTCTCTCCAGAGAGCACCAGCCCTGAAGTCTCTACACCATTTCCACTCTT 1556
QY 1501 CCGCCAGAACCTGAGGGTGTGGGAGAGAAAGTTGTAGGCATGGAAGAGCCAGGCA 1560
DB 1557 CCGCCAGAACCTGAGGGTGTGGGAGAGAAAGTTGTAGGCATGGAAGAGCCAGGCA 1616
QY 1561 GCGCCGAAAGAGAGCTGGGCCACGGCCACGTATAAACCCTCCGAGGGGTCAAGAGTCC 1620
DB 1617 GCGCCGAAAGAGAGCTGGGCCACGGCCACGTATAAACCCTCCGAGGGGTCAAGAGTCC 1676
QY 1621 ATCAGTCTTCTGAGCCCTCTTGAAGCTGAGAGCAGCTCAGAGACCAAGTGAATGCCA 1680
DB 1677 ATCAGTCTTCTGAGCCCTCTTGAAGCTGAGAGCAGCTCAGAGACCAAGTGAATGCCA 1736
QY 1681 GAGGTCTTCTTCCACGAGTCTTCAATGAAGAGCCTCTGACGCCCTTCCACAGCTT 1740
DB 1737 GAGGTCTTCTTCCACGAGTCTTCAATGAAGAGCCTCTGACGCCCTTCCACAGCTT 1796
QY 1741 GCAAGAGCAGAGGAGCCAGCAGTGTGACAGGGGCTCAGCCTGCCCTGAAGTCCCGC 1800
DB 1797 GCAAGAGCAGAGGAGCCAGCAGTGTGACAGGGGCTCAGCCTGCCCTGAAGTCCCGC 1856

QY 1801 CAGTCAGTGGTTACCTCCAGGGCAGTGCCTGTGTGCCCAACCGGAGCCAGCCTTCCAG 1860
DB 1857 CAGTCAGTGGTTACCTCCAGGGCAGTGCCTGTGTGCCCAACCGGAGCCAGCCTTCCAG 1916
QY 1861 GAGCAGAGCAGGGGCGAGGGGCGAGGAGAGCCCTGCATTTCTCTACGCCAGG 1920
DB 1917 GAGCAGAGCAGGGGCGAGGGGCGAGGAGAGCCCTGCATTTCTCTACGCCAGG 1976
QY 1921 TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACGTGAGAGAGGAGGGCGAGCCTGA 1980
DB 1977 TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACGTGAGAGAGGAGGGCGAGCCTGA 2036
RESULT 5
AK094226
LOCUS
DEFINITION
Homo sapiens cDNA FLJ36907 fls, clone BRACE2003800, weakly similar
to MAP kinase phosphatase.
ACCESSION
AK094226
VERSION
AK094226.1 GI:21753246
KEYWORDS
oligo capping; fls (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
1
Tanigami, A., Fujiwara, T., Shibahara, T., Goto, Y., Hirao, M.,
Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K.,
Takahashi, Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2905)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB) (supported by Japan
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
Location/Qualifiers
1..2905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE2003800"
/tissue type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: PME18FL3"
646..2187
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04314.1"
/db_xref="GI:21753247"
/translation="MAFPLSPASSPCTGLVPLWSDTQVLYLDGDFSVTSGQSR
IFKPISTQMTWATLVHQAQCEALGSLVPGSALTWASHYQERLNSQCLNWT
MADLESLRPSAEPGSGSEQEOEQAIRAEIWLVDSDLESVTSKEIRQALELRLGL
PLQYRDFIDNOMLLVAQRDRASRIFFPLYLGEWNAANLELQNRVTHILMARE
IDNEYPERFTYHNRLWDESAQLPHWKETHRFIEARAQGTHTLVCKMGVSRSA
TVLAYAKQYECSLQALRHVQELRPIARPNGFLRQLOIYQGLTASRQSHWQKV
GGVSEHHPAPVSTPRPLPPEPEGGGEKYGMEQAPKEPGRPRINLRGVM
RSISLEPSLESTSETSDMEVFVSSHESHEPQLQFPOLARTKGQGVDRGPQA
LKRQSVVTTLQGSAAVNRTOAQEQEGGCGGCGEPICISSTPRFRVYVROASVHDSG

BASE COUNT 591 a 940 c 867 g 507 t
ORIGIN

Query Match 91.2%; Score 1804.8; DB 9; Length 2905;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 2; Indels 83; Gaps 2;

EEGEA"
66 GGACGAGCGGTCCAGCGAAGAGTGAAGTCCAGCGAAGGCGAGAGCTTTGCGGTCTCCG 125
198 GGACGAGCGGTCCAGCGAAGAGTGAAGTCCAGCGAAGGCGAGAGCTTTGCGGTCTCCG 253
126 TGGGCTGTCTTGGAGTGAAGTGAAGGCGAATGATGATGAGAGAGGCGAGTTC 185
254 TGGGCTGTCTTGGAGTGAAGTGAAGGCGAATGATGATGAGAGAGGCGAGTTC 313
186 TGAGCCAACAGAGAGGCGCGAGTGAAGAGAGTCCACGGGACAGACAGACTTCGG 245
314 TGAGCCAACAGAGAGGCGCGAGTGAAGAGAGTCCACGGGACAGACAGACTTCGG 373
246 GCAAGATCCCAAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 305
374 GCAAGATCCCAAGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
306 ACAGCTGTGAGAGCGCGAGAGATGACATCCGCTGGCAGCCCAAGCTGAGGCAACCCGCGC 365
434 ACAGCTGTGAGAGCGCGAGAGATGACATCCGCTGGCAGCCCAAGCTGAGGCAACCCGCGC 493
366 TCCCGGCTCCGCTACCTGTGTGATTTTCTACACGAGAGAGAGAGAGAGAGAGAGAG 425
494 TCCCGGCTCCGCTACCTGTGTGATTTTCTACACGAGAGAGAGAGAGAGAGAGAGAG 553
426 TGAGAGCGTCTCTCTGGGCGTGAATTTCCCTGACAG----- 461
554 TGAGAGCGTCTCTCTGGGCGTGAATTTCCCTGACAGAGGTTGAGGAGAGAGAGAGAG 613
462 -----CAGCT 466
614 GAGGGCAAGAGGTGAGGGAGAGGGGGAGCAATGCTTTCCCTGTCTCCCTGCCAGCT 673
467 CCCCCAGGTGACACCTGGGCTGTGTCTTCCCTCTGAGTGAACACCCAGGTGACTTAG 526
674 CCCCCAGGTGACACCTGGGCTGTGTCTTCCCTCTGAGTGAACACCCAGGTGACTTAG 733
527 ATGAGAGCGGGGCTTCAAGCTGATCTGTGTGGGCAAGCGGATCTTCAAGCCCATCT 586
734 ATGAGAGCGGGGCTTCAAGCTGATCTGTGTGGGCAAGCGGATCTTCAAGCCCATCT 793
587 CCATCCAGACCATGTGGGCAACTCCAGTATTGCAACCAAGCATGTGAGGCACTCTAG 646
794 CCATCCAGACCATGTGGGCAACTCCAGTATTGCAACCAAGCATGTGAGGCACTCTAG 853
647 GCAGCGGCTTGTACCGGGTGGAGTGCCCTCACTGGGGCCAGCCACTACAGAGAGAGAC 706
854 GCAGCGGCTTGTACCGGGTGGAGTGCCCTCACTGGGGCCAGCCACTACAGAGAGAGAC 913
707 TGAATCCCAACAGAGCTGCTTCAATGATGAGCGGCTATGGCCGAGCTGAGTCTCTGC 766
914 TGAATCCCAACAGAGCTGCTTCAATGATGAGCGGCTATGGCCGAGCTGAGTCTCTGC 973
767 GGCCTCCCAAGCGCGAGCTGGCGGGTCTCAGAAACAGAGAGAGATGAGAGCGGATCC 826
974 GGCCTCCCAAGCGCGAGCTGGCGGGTCTCAGAAACAGAGAGAGATGAGAGCGGATCC 1033
827 GTGCTGAGCTGTGAAAAGTGTGATGTCACTGAGAGAGAGTGTCACTTCCAAAGAGA 886
1034 GTGCTGAGCTGTGAAAAGTGTGATGTCACTGAGAGAGAGTGTCACTTCCAAAGAGA 1093
887 TCCGCCAGGCTGTGAGCTGCGCTGGGGCTCCCCCTCCAGACAGTACCGTGACTTCATCG 946
1094 TCCGCCAGGCTGTGAGCTGCGCTGGGGCTCCCCCTCCAGACAGTACCGTGACTTCATCG 1153
947 ACAACCAAGATGCTGTGTGTGGACAGCGGGAGCCGAGCTTCCGCACTTTCCCCCACC 1006

1154 ACAACCAAGATGCTGTGTGTGGACAGCGGGAGCCGAGCTTCCGCACTTTCCCCCACC 1213
1007 TCTACCTGGGCTCAGAGTGAAGCCAGCAAAACCTTGAGAGAGCTGCAGAGAGACAGGCTCA 1066
1214 TCTACCTGGGCTCAGAGTGAAGCCAGCAAAACCTTGAGAGAGCTGCAGAGAGACAGGCTCA 1273
1067 CCCACATCTTGAACATGAGCGCGGAGATTGACAACTTCACTTGAGCGCTTCACTTACC 1126
1274 CCCACATCTTGAACATGAGCGCGGAGATTGACAACTTCACTTGAGCGCTTCACTTACC 1333
1127 ACAATGTGGGCTCTGGATGAGAGTGGGCCAGCTGCTGCCACTGGAAGAGAGAGCGC 1186
1334 ACAATGTGGGCTCTGGATGAGAGTGGGCCAGCTGCTGCCACTGGAAGAGAGAGCGC 1393
1187 ACCGCTTCAATTGAGGCTGCAAGAGACAGGGCACCACGCTGTGTCCACTGCAAGATGG 1246
1394 ACCGCTTCAATTGAGGCTGCAAGAGACAGGGCACCACGCTGTGTCCACTGCAAGATGG 1453
1247 GCGTACGCGGCTCAGCGGCCACAGTGTCTGGCTATGCCATGAAGCAGTACGAATGCAGCC 1306
1454 GCGTACGCGGCTCAGCGGCCACAGTGTCTGGCTATGCCATGAAGCAGTACGAATGCAGCC 1513
1307 TGAGCAGAGCGCTTGGCCACGTGCAAGAGAGCTCCGGGCCATCGCCGCCCAACCTGGCT 1366
1514 TGAGCAGAGCGCTTGGCCACGTGCAAGAGAGCTCCGGGCCATCGCCGCCCAACCTGGCT 1573
1367 TCCTGCGCCAGCTGACAGATCTACAGGGCATCTGACGGGCCAGCCGCAAGAGAGAGAG 1426
1574 TCCTGCGCCAGCTGACAGATCTACAGGGCATCTGACGGGCCAGCCGCAAGAGAGAGAG 1633
1427 GGGAGCAGAAAGTGGGTGGGCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1486
1634 GGGAGCAGAAAGTGGGTGGGCTTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1693
1487 CATTCCCAAGCTTCCGCGAGAACTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAG 1546
1694 CATTCCCAAGCTTCCGCGAGAACTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAG 1753
1547 AAGAGAGCGAGCAGCCCCGAAAGAGAGAGCTGGGCCAGCGCCACGTATAAACCTCCGAG 1606
1754 AAGAGAGCGAGCAGCCCCGAAAGAGAGAGCTGGGCCAGCGCCACGTATAAACCTCCGAG 1813
1607 GGGTCATGAGTCCATCAGTCTTCTGAGAGCCCTCTTGAGAGCTGAGAGAGAGAGAGAG 1666
1814 GGGTCATGAGTCCATCAGTCTTCTGAGAGCCCTCTTGAGAGCTGAGAGAGAGAGAGAG 1873
1667 CCAAGTACATGCGCAGAGGCTTCTCTCCACAGAGTCTTCAATGAAGAGAGCTTGCAGC 1726
1874 CCAAGTACATGCGCAGAGGCTTCTCTCCACAGAGTCTTCAATGAAGAGAGCTTGCAGC 1933
1727 CCTTCCCAAGCTTGAAGAGCCAGAGAGAGGCGCAGAGAGTGAAGAGAGAGAGAGAGAG 1786
1934 CCTTCCCAAGCTTGAAGAGCCAGAGAGAGGCGCAGAGAGTGAAGAGAGAGAGAGAGAG 1993
1787 CCCTGAAGTCCCGCCAGTCAAGTGTATCCCTCCAGGGCAGTGCCTGTGTGCGCAACCGGA 1846
1994 CCCTGAAGTCCCGCCAGTCAAGTGTATCCCTCCAGGGCAGTGCCTGTGTGCGCAACCGGA 2053
1847 CCCAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1906
2054 CCCAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2113
1907 CCTCTACGCCCAAGCTTCCGGAAGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGG 1966
2114 CCTCTACGCCCAAGCTTCCGGAAGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGG 2173
1967 AGGGCAGGCTGA 1980
2174 AGGGCAGGCTGA 2187

AX574785
LOCUS AX574785 2704 bp DNA linear PAT 07-JAN-2003
DEFINITION Sequence 1 from Patent WO0242436.
ACCESSION AX574785
VERSION AX574785.1 GI:27551934
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1
AUTHORS Wei,M.H., Ketchum,K.A., di Francesco,V. and Beasley,E.M.
TITLE Isolated human phosphatase proteins, nucleic acid molecules
encoding human phosphatase proteins, and uses thereof
JOURNAL Patent: WO 0242436-A 1 30-MAY-2002;
PE Corporation (NY) (US)
FEATURES
source
location/Qualifiers
1..2704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 569 a 874 c 794 g 467 t
ORIGIN
Query Match 89.6%; Score 1774; DB 6; Length 2704;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1882; Conservative 0; Mismatches 0; Indels 98; Gaps 1;
QY 1 ATGGCCCTGGTCACTGAGCCGTTCCGCCCCGGGAGCGGCGCTCCACGCCGTGGGG 60
Db 94 ATGGCCCTGGTCACTGAGCCGTTCCGCCCCGGGAGCGGCGCTCCACGCCGTGGGG 153
QY 61 CCCTGGGACCGAGCGGTCCAGCGAAGAGTCCACTCCAGCGAAGCAGAGCTTGGCGTG 120
Db 154 CCCTGGGACCGAGCGGTCCAGCGAAGAGTCCACTCCAGCGAAGCAGAGCTTGGCGTG 213
QY 121 CTCCGTGGGGCTGTCTCTGGGACTGCAAGATGAGGGGACAATGATGATGACAGAGGCC 180
Db 214 CTCCGTGGGGCTGTCTCTGGGACTGCAAGATGAGGGGACAATGATGATGACAGAGGCC 273
QY 181 AGTTCTGAGCCACAGAGAAGGCCCGGAGTGAGAGAGCTCCAGGGGACCAAGACAGAC 240
Db 274 AGTTCTGAGCCACAGAGAAGGCCCGGAGTGAGAGAGCTCCAGGGGACCAAGACAGAC 333
QY 241 TTCGGGCAAGGATCCAGAGTCCCGCAGAGCAGAGAGAGAGGACGACCTGCACTC 300
Db 334 TTCGGGCAAGGATCCAGAGTCCCGCAGAGCAGAGAGAGAGGACGACCTGCACTC 393
QY 301 ATGTACAGCTGTGAGGCGCAGAGATGATCCGCTGGCAGCCAGCTGAGGCAACC 360
Db 394 ATGTACAGCTGTGAGGCGCAGAGATGATCCGCTGGCAGCCAGCTGAGGCAACC 453
QY 361 CGGCTCCCGGCTCCGCTACCTGTGTAATTCTACAGAGAAGAGAGAGTCTGAGC 420
Db 454 CGGCTCCCGGCTCCGCTACCTGTGTAATTCTACAGAGAAGAGAGTCTGAGC 513
QY 421 CAGGATGAGAGCGTCTCTCTGGCGTGAATTTCCTGACAGAGCTCCCGCAGCTGACC 480
Db 514 CAGGATGAGAGCGTCTCTCTGGCGTGAATTTCCTGACAGAGCTCCCGCAGCTGACC 573
QY 481 CTGGGCTGGTCTTGGCCCTCTGAGTGAACCCAGGTGACTTAGATGAGAGCGGGGC 540
Db 574 CTGGGCTGGTCTTGGCCCTCTGAGTGAACCCAGGTGACTTAGATGAGAGCGGGGC 633
QY 541 TTCAGCGTGAAGTCTGGTGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
Db 634 TTCAGCGTGAAGTCTGGTGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATG 693
QY 601 TGGGCCACACTCCAGGATTTGACCAAGCATGTGAGGAGCTCTAGGAGCGGCTTGTGA 660
Db 694 TGGGCCACACTCCAGGATTTGACCAAGCATGTGAGGAGCTCTAGGAGCGGCTTGTGA 753
QY 661 CCGGGTGGAGTCCCTCACTGGGGCAGCCACTACAGAGAGAGACTGAACCTCCGAACAG 720

Db 754 CCGGGTGGAGTCCCTCACTGGGGCAGCCACTACAGAGAGAGACTGAACCTCCGAACAG 813
QY 721 AGTGCCCTCAATGAGTGAACGCGCTATGCGCCGACCTGGAGTCTCTGCGGCTCCAGCGCC 780
Db 814 AGTGCCCTCAATGAGTGAACGCGCTATGCGCCGACCTGGAGTCTCTGCGGCTCCAGCGCC 873
QY 781 GAGCTGGCGGCTCTCAGAACAGAGAGAGATGAGAGCAGCGGATCCGTGTGAGCTGTGG 840
Db 874 GAGCTGGCGGCTCTCAGAACAGAGAGAGATGAGAGCAGCGGATCCGTGTGAGCTGTGG 933
QY 841 AAAGTTGATGATGATGAGTGAACCTGAGAGTCACTTCCAAAGAGATCCGCGAGGCTTG 900
Db 934 AAAGTTGATGATGATGAGTGAACCTGAGAGTCACTTCCAAAGAGATCCGCGAGGCTTG 993
QY 901 GAGCTGGCGGCTCTCAGAACAGAGAGAGTGAACCTGAGAGTCACTTCCAAAGAGATCCG 960
Db 994 GAGCTGGCGGCTCTCAGAACAGAGAGAGTGAACCTGAGAGTCACTTCCAAAGAGATCCG 1053
QY 961 CTGCTGTGGCAGACCGGGGACCGGAGCTCCCGCATCTTCCCGCACCTTACCTGGGCTCA 1020
Db 1054 CTGCTGTGGCAGACCGGGGACCGGAGCTCCCGCATCTTCCCGCACCTTACCTGGGCTCA 1113
QY 1021 GAGTGAACCGCAGAAACCTGAGAGAGTGCAGAGAGAGAGAGAGAGGATCACCACATCTTGAAC 1080
Db 1114 GAGTGAACCGCAGAAACCTGAGAGAGTGCAGAGAGAGAGAGAGAGGATCACCACATCTTGAAC 1173
QY 1081 ATGGCCCGGAGATTTGAACAATCTTCACTTCAAGCGCTTCACTTCAACACATGTGCGCTC 1140
Db 1174 ATGGCCCGGAGATTTGAACAATCTTCACTTCAAGCGCTTCACTTCAACACATGTGCGCTC 1233
QY 1141 TGGGATGAGAGATGGGCCCGCAGCTGTCTGCGCAGCTGGAAGAGAGAGCAGCCGCTTCAATTGAG 1200
Db 1234 TGGGATGAGAGATGGGCCCGCAGCTGTCTGCGCAGCTGGAAGAGAGAGCAGCCGCTTCAATTGAG 1293
QY 1201 GCTGCAAGAGACAGAGGACCCACAGCTGTCTGCTCCACTGCAAGATGGGCTGAGCGGCTCA 1260
Db 1294 GCTGCAAGAGACAGAGGACCCACAGCTGTCTGCTCCACTGCAAGATGGGCTGAGCGGCTCA 1353
QY 1261 GCGGCCACAGTGTGCTGCTATGCTATGCAAGCAGTACGATGAGAGAGAGGAGAGGCTG 1320
Db 1354 GCGGCCACAGTGTGCTGCTATGCTATGCAAGCAGTACGATGAGAGAGAGGAGAGGCTG 1413
QY 1321 CGCCAGTGCAGAGAGCTCCGCGCCATGCGCCGCGCCGCGCCGCGCTTCTGCGCCAGCTG 1380
Db 1414 CGCCAGTGCAGAGAGCTCCGCGCCATGCGCCGCGCCGCGCCGCGCTTCTGCGCCAGCTG 1473
QY 1381 CAGATCTACAGGAGATCTGACGGCCAGCCGCGAGCATGTCTGGAGAGAGAAAGTG 1440
Db 1474 CAGATCTACAGGAGATCTGACGGCCAGCCGCGAGCATGTCTGGAGAGAGAAAGTG 1497
QY 1441 GGTGGGTCTCCCAAGAGAGAGACCCAGCCCTGAAGTCTCTACACCATTCACCTCTT 1500
Db 1498 ----- 1497
QY 1501 CCGCCAGAACCTGAGGGGTGGTGGGAGAGAGAGTTGTAGGATGGAAGAGAGAGAGAGCA 1560
Db 1498 --GCCAGAACCTGAGGGGTGGTGGGAGAGAGAGTTGTAGGATGGAAGAGAGAGAGCA 1555
QY 1561 GCGCCGAAAGAGAGAGCTTGGGCCACGGCCACGTATAAAGCTCCGAGGGGTCAAGAGTCC 1620
Db 1556 GCGCCGAAAGAGAGAGCTTGGGCCACGGCCACGTATAAAGCTCCGAGGGGTCAAGAGTCC 1615
QY 1621 ATCAGTCTTGTGAGAGCTCTTGTGAGCTGAGAGAGCACTCAGAGAGAGAGAGAGAGTCC 1680
Db 1616 ATCAGTCTTGTGAGAGCTCTTGTGAGCTGAGAGAGCACTCAGAGAGAGAGAGAGTCC 1675
QY 1681 GAGGTCTTCTTCCACAGAGTCTTCAATGAAGAGCTTGTGAGAGAGCTTCCCAAGCTT 1740
Db 1676 GAGGTCTTCTTCCACAGAGTCTTCAATGAAGAGCTTGTGAGAGAGCTTCCCAAGCTT 1735
QY 1741 GCAAGGACCAAGGAGGCGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800

Db 1736 GCAAGGACCAAGGAGGAGCCAGAGGTGACAGAGGGGGCTCAGCCTGCCCTGAAGTCCCGC 1795

QY 1801 CAGTCAGTGGTTACCCCTCCAGGGGAGTGCCTGGTGGCCCAACCGGAGCCAGGCTTCCAG 1860

Db 1796 CAGTCAGTGGTTACCCCTCCAGGGGAGTGCCTGGTGGCCCAACCGGAGCCAGGCTTCCAG 1855

QY 1861 GAGCAGGAGCAGGGGAGGGGAGGGGAGAGAGCCCTGCATTCTCTACGCCCCAGG 1920

Db 1856 GAGCAGGAGCAGGGGAGGGGAGGGGAGAGAGCCCTGCATTCTCTACGCCCCAGG 1915

QY 1921 TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGGCCGAGCCTGA 1980

Db 1916 TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGGGCCGAGCCTGA 1975

RESULT 7

AX574791 2704 bp DNA linear PAT 07-JAN-2003

LOCUS AX574791

DEFINITION Sequence 7 from Patent WO0242436.

ACCESSION AX574791

VERSION AX574791.1 GI:27551937

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

1 Wei,M.H., Ketchum,K.A., di Francesco,V. and Beasley,E.M.
Isolated human phosphatase proteins, nucleic acid molecules
encoding human phosphatase proteins, and uses thereof
Patent: WO 0242436-A 7 30-MAY-2002;
PE Corporation (NY) (US)
Location/Qualifiers
1.2704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 569 a 874 C 794 G 467 t

ORIGIN

Query Match 89.6%; Score 1774; DB 6; Length 2704;
Best Local Similarity 95.1%; Pred.No. 0;
Matches 1882; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

QY 1 ATGCCCTGGTCAACAGTGAAGCCGTTCCGCCCGGAGCGGCGCTCCAGCGCCCGTGGGG 60

Db 94 ATGCCCTGGTCAACAGTGAAGCCGTTCCGCCCGGAGCGGCGCTCCAGCGCCCGTGGGG 153

QY 61 CCCTGGGAGCAGGCGGCTCCAGCGGAGAGAGTGAAGTCACTCCAGCGGAGGAGAGCTTTGGGGT 120

Db 154 CCCTGGGAGCAGGCGGCTCCAGCGGAGAGAGTGAAGTCACTCCAGCGGAGGAGAGCTTTGGGGT 213

QY 121 CTCCGTGGGGCTGTCTCTGGGAGTGCAGGATGAGGGGGAACAATGATGATGACAGAGGCC 180

Db 214 CTCCGTGGGGCTGTCTCTGGGAGTGCAGGATGAGGGGGAACAATGATGATGACAGAGGCC 273

QY 181 AGTTCTGAGCCCAACAGAGAAAGCCCGAGTGAAGAGAGAGTCTCCAGCGGGAGCCAGACAGAC 240

Db 274 AGTTCTGAGCCCAACAGAGAAAGCCCGAGTGAAGAGAGAGTCTCCAGCGGGAGCCAGACAGAC 333

QY 241 TTCCGGCAAGGATCCACAGAGTCCCAAGAGCAGAGGAGAGAGGAGGAGCCTGCACCTC 300

Db 334 TTCCGGCAAGGATCCACAGAGTCCCAAGAGCAGAGGAGAGAGGAGGAGCCTGCACCTC 393

QY 301 ATGTAACAGCTGTGAGGCGCAGAGATGACATCCGCTGGCAGGCCAGCTGAGGACACC 360

Db 394 ATGTAACAGCTGTGAGGCGCAGAGATGACATCCGCTGGCAGGCCAGCTGAGGACACC 453

QY 361 CGGCTCCCGGCTCCGCTACCTGTGTAGTTCTACAGAGAGAGAGAGAGGTTGTAGC 420

Db 454 CGGCTCCCGGCTCCGCTACCTGTGTAGTTCTACAGAGAGAGAGAGAGGTTGTAGC 513

QY 421 CAGATGAGACGGTCTCTCTGGGGCTGGAATTTCCCTGACAGCAGCTCCCGCAGCTGACC 480

Db 514 CAGGATGAGACGGTCTCCTCGGGCGTGGATTTCCCTGACAGACGCTCCCCAGCTGACC 573

QY 481 CTGGGCTGGTCTTGCCCTCTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGC 540

Db 574 CTGGGCTGGTCTTGCCCTCTGAGTGACACCCAGGTGTACTTAGATGAGACGGGGC 633

QY 541 TTCAAGCGTGAAGTCTGTGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 600

Db 634 TTCAAGCGTGAAGTCTGTGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG 693

QY 601 TGGGCGACACTCCAGGTATTGCAACCAAGCATGTGAGGCAAGCTTAGGCGAGCGGCTTGT 660

Db 694 TGGGCGACACTCCAGGTATTGCAACCAAGCATGTGAGGCAAGCTTAGGCGAGCGGCTTGT 753

QY 661 CCGGGTGGCAGTGCCTCCTACCTGGGGCCAGCCACTTACAGAGAGAGACTGAATCCGAACAG 720

Db 754 CCGGGTGGCAGTGCCTCCTACCTGGGGCCAGCCACTTACAGAGAGAGACTGAATCCGAACAG 813

QY 721 AGCTGCCCTCAATGAGTGAAGCGGCTATGGCCGAGCTGAGTCTTGCGGCTCCAGCGCC 780

Db 814 AGCTGCCCTCAATGAGTGAAGCGGCTATGGCCGAGCTGAGTCTTGCGGCTCCAGCGCC 873

QY 781 GAGCCTGGCGGGTCTCTCAGAAACAGAGCAGATGAGAGCAGGCGATCCGTGCTGAGCTGTGG 840

Db 874 GAGCCTGGCGGGTCTCTCAGAAACAGAGCAGATGAGAGCAGGCGATCCGTGCTGAGCTGTGG 933

QY 841 AAAGTTGATGATGTCAGTGAAGTGAAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 900

Db 934 AAAGTTGATGATGTCAGTGAAGTGAAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 993

QY 901 GAGCTGGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAAACAGATGCTG 960

Db 994 GAGCTGGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAAACAGATGCTG 1053

QY 961 CTGCTGTGGCACAGCGGGAGCCGAGCTCCCGCATCTTCCCCACCTCTACCTGGGCTCA 1020

Db 1054 CTGCTGTGGCACAGCGGGAGCCGAGCTCCCGCATCTTCCCCACCTCTACCTGGGCTCA 1113

QY 1021 GAGTGAACCGCAGCAACCTGGAGAGAGTGCAGAGAGAAACAGGTCACCCACATCTTGAAC 1080

Db 1114 GAGTGAACCGCAGCAACCTGGAGAGAGTGCAGAGAGAAACAGGTCACCCACATCTTGAAC 1173

QY 1081 ATGGCCCGGAGATTGACAACTTTCAACCTGAAGCGCTTCACTTACCACAATGTGCGCTC 1140

Db 1174 ATGGCCCGGAGATTGACAACTTTCAACCTGAAGCGCTTCACTTACCACAATGTGCGCTC 1233

QY 1141 TGGGATGAGAGATCGGCCAGCTGCTGCCGCACTGGAAGAGAGACGACCCGCTTCAATTGAG 1200

Db 1234 TGGGATGAGAGATCGGCCAGCTGCTGCCGCACTGGAAGAGAGACGACCCGCTTCAATTGAG 1293

QY 1201 GCTGCAAGACACAGGGGACCCACGCTGCTGTGCTCACTGCAAGATGAGGCGCTGAGCGGCTCA 1260

Db 1294 GCTGCAAGACACAGGGGACCCACGCTGCTGTGCTCACTGCAAGATGAGGCGCTGAGCGGCTCA 1353

QY 1261 GCGGCAAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACAGCCTGAGCAGGCGCTG 1320

Db 1354 GCGGCAAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACAGCCTGAGCAGGCGCTG 1413

QY 1321 CGCCAGTGAAGAGTCCGGCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1380

Db 1414 CGCCAGTGAAGAGTCCGGCCATGCGCCGCCCAACCTGGCTTCTGCGCCAGCTG 1473

QY 1381 CAGATCTACAGGAGCATCTGACGGCCAGCGGCAAGCCATGTCTGGAGAGAGAAAGTG 1440

Db 1474 CAGATCTACAGGAGCATCTGACG----- 1497

QY 1441 GTGGGGTCTCCCAAGAGAGCAACCGCCCTGAAGTCTCTACACCATTTCCCACTCTT 1500

Db 1498 ----- 1497

QY 1501 CCGCAGAACTGAGGTTGTGGGAGAGAGAGGTTGAGCATGGAAGAGCCAGGCA 1560

QY 841 AAAGTGTGATGTAGTACTGAGAGTGTCACTTCCAAGAGATCCGCGAGCTCTG 900
DB 841 AAAGTGTGATGTAGTACTGAGAGTGTCACTTCCAAGAGATCCGCGAGCTCTG 900
QY 901 GAGCTGCGCTGGGCTCCCTCCAGACGTACCGTGACTTTCATCGAACACAGATGCTG 960
DB 901 GAGCTGCGCTGGGCTCCCTCCAGACGTACCGTGACTTTCATCGAACACAGATGCTG 960
QY 961 CTGCTGTGGCAGACGGGACCGAGCCTCCCGCATCTTCCCGCAGCTTACCTGGGCTCA 1020
DB 961 CTGCTGTGGCAGACGGGACCGAGCCTCCCGCATCTTCCCGCAGCTTACCTGGGCTCA 1020
QY 1021 GAGTGAACGCAACCACTGAGAGAGCTGCAGAGGAACAGGTCACCCACATCTTGAAC 1080
DB 1021 GAGTGAACGCAACCACTGAGAGAGCTGCAGAGGAACAGGTCACCCACATCTTGAAC 1080
QY 1081 ATGCCCCGGGAGATTGACAATTCTACCTGAGCGCTTCACTACCAATGTGCGCTC 1140
DB 1081 ATGCCCCGGGAGATTGACAATTCTACCTGAGCGCTTCACTACCAATGTGCGCTC 1140
QY 1141 TGGGATGAGGAGTGGGCGCCAGCTGCTGCGCACTGGAAGAGACCGACCGCTTCAATTAG 1200
DB 1141 TGGGATGAGGAGTGGGCGCCAGCTGCTGCGCACTGGAAGAGACCGACCGCTTCAATTAG 1200
QY 1201 GCTGCAAGAGCAGAGGGGCAACCGTGTGCTGCTCACTGCAAGATGGGCGTCA 1260
DB 1201 GCTGCAAGAGCAGAGGGGCAACCGTGTGCTGCTCACTGCAAGATGGGCGTCA 1260
QY 1261 GCGGCCACAGTGTGCTGCTATGCTATGCAAGAGTACGATGAGAGCGAGCGCTG 1320
DB 1261 GCGGCCACAGTGTGCTGCTATGCTATGCAAGAGTACGATGAGAGCGAGCGCTG 1320
QY 1321 CGCCACGTGAGAGAGCTCGGCGCCATCGCCCGCCCAACCTGCTTCTGCGCAGCTG 1380
DB 1321 CGCCACGTGAGAGAGCTCGGCGCCATCGCCCGCCCAACCTGCTTCTGCGCAGCTG 1380
QY 1381 CAGATCTACAGGCGATCTGACG----- 1404
DB 1381 CAGATCTACAGGCGATCTGACG----- 1404
QY 1441 GGTGGGGTCTCCCGAGAGGCAACCCAGCCCTGAAGTCTTACACCAATCCACCTCT 1500
DB 1441 ----- 1404
QY 1501 CCGCCAGAACCTGAGGGTGGTGGGAGAGAGAGTTGTAGGCATGGAAGAGCCAGGCA 1560
DB 1405 --GCCAGAACCTGAGGGTGGTGGGAGAGAGAGTTGTAGGCATGGAAGAGCCAGGCA 1462
QY 1561 GCCCCGAAAGAGAGAGCTGGGCCACGGCCACGTAATAAACCCTCCGAGGGGTCTAGAGTCC 1620
DB 1463 GCCCCGAAAGAGAGAGCTGGGCCACGGCCACGTAATAAACCCTCCGAGGGGTCTAGAGTCC 1522
QY 1621 ATCAGTCTTCTGAGAGCCCTCTTGGAGCTGGAAGACCACTCAGAGACCAAGTACATGCCA 1680
DB 1523 ATCAGTCTTCTGAGAGCCCTCTTGGAGCTGGAAGACCACTCAGAGACCAAGTACATGCCA 1582
QY 1681 GAGGTCTTCTTCCACAGAGTCTTCAATGAAGAGCTCTGACGCCCTTCCACAGCTT 1740
DB 1583 GAGGTCTTCTTCCACAGAGTCTTCAATGAAGAGAGCTCTGACGCCCTTCCACAGCTT 1642
QY 1741 GCAAGGACCAAGGGAGGCCAGAGTGAACAGGGGGCTCAGCCTGCCCTGAAGTCCGC 1800
DB 1643 GCAAGGACCAAGGGAGGCCAGAGTGAACAGGGGGCTCAGCCTGCCCTGAAGTCCGC 1702
QY 1801 CAGTCACTGTTACCTCTCAGAGGAGTGCCTGTGGCCAACCGGACCGAGGCTTCCAG 1860
DB 1703 CAGTCACTGTTACCTCTCAGAGGAGTGCCTGTGGCCAACCGGACCGAGGCTTCCAG 1762
QY 1861 GAGCAGAGCAGGGGAGGGGAGGGGAGGGGAGAGCCCTGATTTCTTCTAGCCCGAG 1920
DB 1763 GAGCAGAGCAGGGGAGGGGAGGGGAGGGGAGAGCCCTGATTTCTTCTAGCCCGAG 1822
QY 1921 TTCCGGAAGGTGTGAGACAGGCCAGCGTGAAGTGAAGTGAAGAGAGGGCGAGGCTGA 1980

DB 1823 TTCCGGAAGGTGTGAGACAGGCCAGCGTGATGACAGTGAAGAGAGGGCGAGGCTGA 1882
RESULT 9
AX398774
LOCUS AX398774 2322 bp DNA linear PAT 27-MAY-2002
DEFINITION Sequence 10 from Patent WO0220732.
ACCESSION AX398774
VERSION AX398774.1 GI:21261307
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Liou, J.R.
TITLE Regulation of human map kinase phosphatase-like enzyme
JOURNAL Patent: WO 0220732-A 10 14-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
source location/Qualifiers
1..2322
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 467 a 751 c 718 g 386 t
ORIGIN
Query Match 85.1%; Score 1684; DB 6; Length 2322;
Best Local Similarity 95.0%; Pred. No. 2.6e-307;
Matches 1852; Conservative 0; Mismatches 75; Indels 23; Gaps 10;
QY 39 CGGCGCTTCCACGCGCGTGGGCGCTTGGAGACAGCGCGTCCAGGAGAGTGAAGTCCA 98
DB 147 CGGCGTCACTTGGGCGTGTCTCTCGGAGAGAGTGTCCAGGAGAGTGAAGTCCA 206
QY 99 GCGAAGCAGAGCTTTGCGGTGCTCC-GTGGGGTGTCTGGAAGTGAAGGGG 157
DB 207 GCGA----AGAGCTTTGCGGTGCTCTGTGGGGTGTCTGGAAGTGAAGGGG 262
QY 158 ACAATGATGATGACAGAGAGGCCAGTCTGAGCCACAAGAGAGGCCCGAGTGAAGAG 217
DB 263 ACAATGATGATGACAGAGAGGCCAGTCTGAGCCACAAGAGAGGCCCGAGTGAAGAG 322
QY 218 AGCTCCAGCGGG--ACCAACAGACTT--CGGGCAAGATCCCGAGTCCCGAGAAGCA 272
DB 323 AGCTCCAGCGGGGAACCCAGACAGACTTTCCGTGCAAGATCCCGAGTCCCGAGAAGCA 382
QY 273 GGAGGACAGAGGACGACCTGCACTCATGTGTAAGCTGTAAGGCCGAGAGATGACAT 332
DB 383 GGAGGACAGAGGACGACCTGCACTCATGTGTAAGGGCTGTAAGGCCGAGAGATGACAT 442
QY 333 CCGGCTGGCA--GCCAGCTGGAGGCAACCCCGGCTCCCGGCTCCCGCTACCTGC----- 385
DB 443 CCGACTTGAAGCCCAAGCTGGAGGCAACCCCGGCTCCCGGATCCGATACCTTGCT 502
QY 386 TGGTAGTTTCTACAGAGAA-GAAGAGTCTGAG-CCAGATGAGACGGTCTCTGGG 443
DB 503 TGGTAGTTTCTACAGAGAGGAAGAGTCTGAGGCCAGATGAGACGGTCTCTGGG 562
QY 444 -CGTGATTTCCCTGACAGAGCTCCCGAGCTGACCTGGGCTGGTCTGCCCTCT 502
DB 563 ACGTGATTTCCCTGACAGAGCTCCCGAGCTGACCTGGGCTGGTCTGCCCTCT 622
QY 503 GGAGTGAACCCAGGTGTACTTATGATGAGAGCGGGCTTACGCTGAGCTGTGGGC 562
DB 623 GGAGTGAACCCAGGTGTACTTATGAGAGCGGGCTTACGCTGAGCTGTGGGC 682
QY 563 AAAGCGGATTTCAAGCCCATCTCCATCCAGACATGTGGCCACACTCCAGTATTGC 622
DB 683 AAAGCGGATTTCAAGCCCATCTCCATCCAGACATGTGGCCACACTCCAGTATTGC 742
QY 623 ACCAAGATGTGAGCAGCTCTAGGACGGGCTTGTACCGGAGTGGAGTGCCTCACCT 682

Db 743 ACCAAGCATGTGAGGAGCTCTGAGGACGGGCTTGTACCGGGTGAGTGCCCTCACCCT 802
QY 683 GGGCCAGCACTACAGAGAGAGTGAATCCGAAACAGAGCTGCTCAATGAGTGACGG 742
Db 803 GGGCCAGCACTACAGAGAGAGTGAATCCGAAACAGAGCTGCTCAATGAGTGACGG 862
QY 743 CTATGGCCGACTGAGTCTCTGCGGCTCCAGCGCCGAGCCTGGCGGGTCTCAGAAC 802
Db 863 CTATGGCCGACTGAGTCTCTGCGGCTCCAGCGCCGAGCCTGGCGGGTCTCAGAAC 922
QY 803 AGGAGCAGATGAGAGCAGGCGATCCGTGCTGAGCTGTGAAAAGTGTGATGTGAGTACC 862
Db 923 AGGAGCAGATGAGAGCAGGCGATCCGTGCTGAGCTGTGAAAAGTGTGATGTGAGTACC 982
QY 863 TGGAGAGTGTGCTTCCAGAGAGATCCGCGCAGGCTGTGAGCTGCGCTGGGGCTCCCCC 922
Db 983 TGGAGAGTGTGCTTCCAGAGAGATCCGCGCAGGCTGTGAGCTGCGCTGGGGCTCCCCC 1042
QY 923 TCCAGCAGATCCGCTGACTTCAATCAACAGATGTGCTGCTGCTGGGCAAGCGGGAGCC 982
Db 1043 TCCAGCAGATCCGCTGACTTCAATCAACAGATGTGCTGCTGCTGGGCAAGCGGGAGCC 1102
QY 983 GAGCCTCCCGCATCTTCCCCCACCCTCTAACCCTGCTCAGAGTGGAACGCCAAACCTGG 1042
Db 1103 GAGCCTCCCGCATCTTCCCCCACCCTCTAACCCTGCTCAGAGTGGAACGCCAAACCTGG 1162
QY 1043 AGGAGCTGAGAGGAAACAGGGGTACCCCAATCTTGAACATGGCCCCGGAGATTGACACT 1102
Db 1163 AGGAGCTGAGAGGAAACAGGGGTACCCCAATCTTGAACATGGCCCCGGAGATTGACACT 1219
QY 1103 TCTACCTGAGCGGCTTCACTTACCAATGTGCGGCTCTGGGATGAGGATGGGCCAGC 1162
Db 1220 TCTACCTGAGCGGCTTCACTTACCAATGTGCGGCTCTGGGATGAGGATGGGCCAGC 1279
QY 1163 TGCTGCGCAGTGAAGAGAGAGCGACCGCTTCAATGAGGCTGCAAGAGCAGAGGGCACCC 1222
Db 1280 TGCTGCGCAGTGAAGAGAGAGCGACCGCTTCAATGAGGCTGCAAGAGCAGAGGGCACCC 1339
QY 1223 ACGTGTGTCTCACTGCAAGATGGGGGTGAGCGGCTCAGCGGCCACAGTGTGCTGCTATG 1282
Db 1340 ACGTGTGTCTCACTGCAAGATGGGGGTGAGCGGCTCAGCGGCCACAGTGTGCTGCTATG 1399
QY 1283 CCATGAAGCAGTACGAATGAGAGCTGAGAGCGGCTGCGCCACAGTGACAGAGCTCCGGC 1342
Db 1400 CCATGAAGCAGTACGAATGAGAGCTGAGAGCGGCTGCGCCACAGTGACAGAGCTCCGGC 1459
QY 1343 CCATCGCCCGCCCAACCTGGCTTCTGCGCAGCTGAGATCTTACAGGGCATCTCTGA 1402
Db 1460 CCATCGCCCGCCCAACCTGGCTTCTGCGCAGCTGAGATCTTACAGGGCATCTCTGA 1519
QY 1403 CGGCGAGCGCCAGAGCCATGTCTGGGAGCAGAAAGTGGGTGGGTCTCCCAAGAGAGC 1462
Db 1520 CGGCGAGCGCCAGAGCCATGTCTGGGAGCAGAAAGTGGGTGGGTCTCCCAAGAGAGC 1579
QY 1463 ACCAGGCCCTGAAGTCTCTACACCATTTCCCACTTCCGCCAGAACCTGAGGGTGTG 1522
Db 1580 ACCAGGCCCTGAAGTCTCTACACCATTTCCCACTTCCGCCAGAACCTGAGGGTGTG 1639
QY 1523 GGGAGGAGAGGTTGTAGGAGTGAAGAGAGCCAGGCGAGCCCGGAAAGAGAGCCTGGGC 1582
Db 1640 GGGAGGAGAGGTTGTAGGAGTGAAGAGAGCCAGGCGAGCCCGGAAAGAGAGCCTGGGC 1699
QY 1583 CAGGCGCACGTATAAACCCTCCGAGGGGTATGAGGTCCATCATCTTCTTGAGCCCTCCT 1642
Db 1700 CAGGCGCACGTATAAACCCTCCGAGGGGTATGAGGTCCATCATCTTCTTGAGCCCTCCT 1759
QY 1643 TGGAGCTGAGAGCAGCTCAGAGACAGTACATGCGCAGAGGCTTCTTCCACAGAGT 1702
Db 1760 TGGAGCTGAGAGCAGCTCAGAGACAGTACATGCGCAGAGGCTTCTTCCACAGAGT 1819
QY 1703 CTTACATGAAGAGCTCTGAGCCCTTCCACAGCTTGAAGGAGCAAGGAGGCGCAGC 1762

Db 1820 CTTACATGAAGAGCTCTGAGCCCTTCCACAGCTTGCAAGAGCAAGGAGGCCAGC 1879
QY 1763 AGGTGACAGGGGGCTCAGCCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1822
Db 1880 AGGTGACAGGGGGCTCAGCCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1939
QY 1823 GCAGTGCCTGTGTGGCCCAACCGGACCCAGGCTTCCAGAGAGCAGAGCAGGAGGAGGAG 1882
Db 1940 GCAGTGCCTGTGTGGCCCAACCGGACCCAGGCTTCCAGAGAGCAGAGCAGGAGGAGGAG 1999
QY 1883 AGGGCAGGAGAGCCCTGATTTCTCTACGCCCAAGTTCGGAAAGTGTGAGACAGG 1942
Db 2000 AGGGCAGGAGAGCCCTGATTTCTCTACGCCCAAGTTCGGAAAGTGTGAGACAGG 2059
QY 1943 CCAGCTGATGACAGTGAAGAGAGGCGG 1972
Db 2060 CCAGCTGATGACAGTGAAGAGAGGCGG 2089

RESULT 10
AX451362 1416 bp DNA linear PAT 03-JUL-2002
LOCUS Sequence 20 from Patent WO0224740.
DEFINITION AX451362
ACCESSION AX451362
VERSION AX451362.1 GI:21698396
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 Luche, R.M. and Wei, B.
AUTHORS Dsp-15 dual-specificity phosphatase
TITLE Patent: WO 0224740-A 20 28-MAR-2002;
JOURNAL Ceplyr, Inc. (US)

FEATURES
source location/Qualifiers
1..1416
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

BASE COUNT 288 a 450 c 449 g 229 t
ORIGIN

Query Match 71.2%; Score 1409; DB 6; Length 1416;
Best Local Similarity 100.0%; Pred. No. 1.9e-255;
Matches 1409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCCCTGTGTCAAGTGAAGCCCTTCCGCGGAGAGGCGGCTCCACGCGCTGGGG 60
Db 1 ATGGCCCTGTGTCAAGTGAAGCCCTTCCGCGGAGAGGCGGCTCCACGCGCTGGGG 60
QY 61 CCCTGGAGCAGGCGGTCCAGCGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 120
Db 61 CCCTGGAGCAGGCGGTCCAGCGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 120
QY 121 CTCCTGGGGTGTCTCTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
Db 121 CTCCTGGGGTGTCTCTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 180
QY 181 AGTTCTGAGCCAAAGAGAGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 240
Db 181 AGTTCTGAGCCAAAGAGAGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 240
QY 241 TTCGGGCAAGATCCAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 TTCGGGCAAGATCCAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 ATGCTACAGCTGTGAGGCGCGAGATGACATCCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 ATGCTACAGCTGTGAGGCGCGAGATGACATCCGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 CGGCTCCCGGCTCCGCTACCTGTGTGAGTGTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

Db 361 CGGCTCCCGCTCCGCTACCTGTGTAGTTTCTACAGAGAGAGAGAGTCTGACC 420
QY 421 CAGGATGAGACGTCCTCTCTGGCGGTGATTTCCCTGACAGACAGCTCCCCAGCTGACC 480
Db 421 CAGGATGAGACGTCCTCTCTGGCGGTGATTTCCCTGACAGACAGCTCCCCAGCTGACC 480
QY 481 CTGGGCTGTGTTGCTCCCTCTGAGAGTACACCCAGGTGTACTTAGTAGAGACGGGGG 540
Db 481 CTGGGCTGTGTTGCTCCCTCTGAGAGTACACCCAGGTGTACTTAGTAGAGACGGGGG 540
QY 541 TTCAGCGTGAAGTCTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
Db 541 TTCAGCGTGAAGTCTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
QY 601 TGGGCCACACTCCAGGTATTGCAACCAAGCATGTAGGCGAGCTTAGGCGAGCGCTTGTA 660
Db 601 TGGGCCACACTCCAGGTATTGCAACCAAGCATGTAGGCGAGCTTAGGCGAGCGCTTGTA 660
QY 661 CCGGCTGGCAGTGCCTTCACTGGGCGCAGCCACTTACAGAGAGAGTGAAGTCCGAACAG 720
Db 661 CCGGCTGGCAGTGCCTTCACTGGGCGCAGCCACTTACAGAGAGAGTGAAGTCCGAACAG 720
QY 721 AGCTGCTCAATGATGATGAGCGCTATGCGCCGACTGGAGTCTCTGCGGCTTCCAGCGCC 780
Db 721 AGCTGCTCAATGATGATGAGCGCTATGCGCCGACTGGAGTCTCTGCGGCTTCCAGCGCC 780
QY 781 GAGCTGGCGGGTCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 GAGCTGGCGGGTCTCTCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 AAAGTGTGATGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 900
Db 841 AAAGTGTGATGTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 900
QY 901 GAGCTGGCGCTGGGCTCCCTCCAGAGTACCGTGAATTCATGCAACAGAGATGCTG 960
Db 901 GAGCTGGCGCTGGGCTCCCTCCAGAGTACCGTGAATTCATGCAACAGAGATGCTG 960
QY 961 CTGCTGTGGGACAGCGGGAACGAGCTCCCGCATCTTCCCGCATCTTACCTGGGCTCA 1020
Db 961 CTGCTGTGGGACAGCGGGAACGAGCTCCCGCATCTTCCCGCATCTTACCTGGGCTCA 1020
QY 1021 GAGTGGAAACGACGAAACCTGGAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 GAGTGGAAACGACGAAACCTGGAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 ATGGCCCGGGAGATTGACAACTTTTACCCTGAGGCTTCACTTACCAATGTGCGCTC 1140
Db 1081 ATGGCCCGGGAGATTGACAACTTTTACCCTGAGGCTTCACTTACCAATGTGCGCTC 1140
QY 1141 TGGGATGAGGAGTGGGCGGCGAGCTGCGGCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 TGGGATGAGGAGTGGGCGGCGAGCTGCGGCACTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GCTGCAAGAGACAGAGGACCCACGCTGTGTCTCACTGCAAGATGGGCTCAGCGCTCA 1260
Db 1201 GCTGCAAGAGACAGAGGACCCACGCTGTGTCTCACTGCAAGATGGGCTCAGCGCTCA 1260
QY 1261 GCGGCAAGTGTGCTGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 GCGGCAAGTGTGCTGCTATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 CGGCAAGTGAAGAGTCCGCGCCATCGCGCCCAACCTGCTTCTGCGCAGCTG 1380
Db 1321 CGGCAAGTGAAGAGTCCGCGCCATCGCGCCCAACCTGCTTCTGCGCAGCTG 1380
QY 1381 CAGATCTACCAAGGAGTCTGAGCGCCAG 1409
Db 1381 CAGATCTACCAAGGAGTCTGAGCGCCAG 1409

RESULT 11
AX574787

LOCUS AX574787 2540 bp DNA linear PAT 07-JAN-2003
DEFINITION Sequence 3 from Patent WO0242436.
ACCESSION AX574787
VERSION AX574787.1 GI:27551936
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Wei, M.H., Ketchum, K.A., di Francesco, V. and Beasley, E.M.
TITLE Isolated human phosphatase proteins, nucleic acid molecules
JOURNAL encoding human phosphatase proteins, and uses thereof
Patent: WO 0242436-A 3 30-MAY-2002;
PE Corporation (NY) (US)
FEATURES
source 1.2540
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 552 a 809 c 739 g 440 t
ORIGIN
Query Match 68.0%; Score 1345.6; DB 6; Length 2540;
Best Local Similarity 85.1%; Pred. No. 1,5e-243;
Matches 1689; Conservative 0; Mismatches 4; Indels 291; Gaps 5;
QY 1 ATGGCCCTGTGTCACAGTGAAGCGTTGCGCCCGGCGAGCGGCGCTCCAGCGCCGTGGG 60
Db 88 ATGGCCCTGTGTCACAGTGAAGCGTTGCGCCCGGCGAGCGGCGCTCCAGCGCCGTGGG 147
QY 61 CCCTGGGACCAAGCGGCTCCAGCGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 120
Db 148 CCCTGGGACCAAGCGGCTCCAGCGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 207
QY 121 CTCCGTGGGCTGTCTCTGGGACTGCAGATGAGAGGAGCAATGATGACAGAGAGGCC 180
Db 208 CTCCGTGGGCTGTCTCTGGGACTGCAGATGAGAGGAGCAATGATGACAGAGAGGCC 267
QY 181 AGTTCTGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 268 AGTTCTGAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 327
QY 241 TTGGGCAAGAGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 328 TTGGGCAAGAGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
QY 301 ATGTACAGCTGCTGAGCGCGAGAGTACATCCGCTGGGAGAGAGAGAGAGAGAGAGAGAG 360
Db 388 ATGTACAGCTGCTGAGCGCGAGAGTACATCCGCTGGGAGAGAGAGAGAGAGAGAGAGAG 447
QY 361 CGGCTCCCGGCTCCGCTACTGCTGTGATTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 448 CGGCTCCCGGCTCCGCTACTGCTGTGATTCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 507
QY 421 CAGATGAGACGCTCCTCTGGGCGGTGATTTCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 508 CAGATGAGACGCTCCTCTGGGCGGTGATTTCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 567
QY 481 CTGGGCTGTGTTGCTCCCTCTGAGTGAACCCAGGTGTACTTAGTAGAGAGAGAGAGAGAGAG 540
Db 568 CTGGGCTGTGTTGCTCCCTCTGAGTGAACCCAGGTGTACTTAGTAGAGAGAGAGAGAGAGAG 627
QY 541 TTCAGCGTGAAGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
Db 628 TTCAGCGTGAAGTGTGGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATG 687
QY 601 TGGGCCACTCCAGTATTGCAACCAAGCATGTGAGGAGCTTAGGAGCGGCTTGTA 660
Db 688 T----- 688
QY 661 CCGGCTGGCAGTGCCTTCACTGGGCGCAGCCACTACAGAGAGAGAGAGAGAGAGAGAGAGAG 720

Db 689 ----- 688
QY 721 AGCTGCCTCAATGATGAGCGGCTATGCGCGACCTGAGTCTTGCGGCTCCAGCGCC 780
Db 689 ----- 688
QY 781 GAGCCTGCGGCTCTCAGAACAGAGAGATGAGCAGCGCATCCGTGCTGAGTGTGG 840
Db 689 ----- 688
QY 841 AAAGTTGATGTAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 900
Db 739 AAAGTTGATGTAGTACCTGAGAGTGTCACTTCCAAAGAGATCCGCGAGGCTCTG 798
QY 901 GAGCTGCGGCTGCGGCTCCCGCTCCAGAGTACCGTGAATTCATCGACACAGATGCTG 960
Db 799 GAGCTGCGGCTGCGGCTCCCGCTCCAGAGTACCGTGAATTCATCGACACAGATGCTG 858
QY 961 CTGCTGTGGCAGACGCGGAGCCGAGCTCCCGCATCTTCCCGCACCTCTACCTGGGCTCA 1020
Db 859 CTGCTGTGGCAGACGCGGAGCCGAGCTCCCGCATCTTCCCGCACCTCTACCTGGGCTCA 918
QY 1021 GAGTGAACCGCAGCAAACTTGAGAGAGTGCAGAGAAACAGAGTCAACCATCTTGAAC 1080
Db 919 GAGTGAACCGCAGCAAACTTGAGAGAGTGCAGAGAAACAGAGTCAACCATCTTGAAC 978
QY 1081 ATGGCCCGGAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCAATGTCGCGCTC 1140
Db 979 ATGGCCCGGAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCAATGTCGCGCTC 1038
QY 1141 TGGGATGAGGAGTGGCCCGCAGCTGTGCGGCACTGGAAGAGACGACCGCTTCAATTGAG 1200
Db 1039 TGGGATGAGGAGTGGCCCGCAGCTGTGCGGCACTGGAAGAGACGACCGCTTCAATTGAG 1098
QY 1201 GCTGCAAGAGCAGAGGCGCACCCAGCTGTGCTGCTCACTGCAAGATGGCGCTCAGCGCTCA 1260
Db 1099 GCTGCAAGAGCAGAGGCGCACCCAGCTGTGCTGCTCACTGCAAGATGGCGCTCAGCGCTCA 1158
QY 1261 GCGGCCACAGTGTGCTGCTATGCAAGCAGTACGATGCAAGCCTGAGCAGCGGCTG 1320
Db 1159 GCGGCCACAGTGTGCTGCTATGCAAGCAGTACGATGCAAGCCTGAGCAGCGGCTG 1218
QY 1321 CGCCACGTGAGAGCTCGGCGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1380
Db 1219 CGCCACGTGAGAGCTCGGCGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1278
QY 1381 CAGATCTACAGGGGATCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1440
Db 1279 CAGATCTACAGGGGATCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1302
QY 1441 GGTGGGTCTCCCCAGAGAGCAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1500
Db 1303 ----- 1302
QY 1501 CCGCCAGAACCTGAGGGTGTGGGAGAGAGGTTGTAGGCATGGAAGAGCCAGGCA 1560
Db 1303 ----- 1302
QY 1561 GCGCCGAAAGAGAGAGCTGGG--CCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1618
Db 1361 GCGCCGAAAGAGAGAGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1420
QY 1619 CCATCAGTCTTCTGAGAGCGGCTCTT--GAGCTGAGAGCAGCTCAG--AGACCAAGTACAT 1676
Db 1421 CCATCAGTCTTCTGAGAGCGGCTCTTGGGAGCTGAGAGCAGCTCAGTAGACCAAGTACAT 1480
QY 1677 GCGCAGAGTCTTCTTCTTCCACAGAGTCTTACATGAAGAGCCTTGCAGCGGCTTCCACA 1736
Db 1481 GCGCAGAGTCTTCTTCTTCCACAGAGTCTTACATGAAGAGCCTTGCAGCGGCTTCCACA 1540
QY 1737 GCTTGAAGAGCAGAGGAGGCGCAGAGTGTGACAGGCGGCTCAGCGCTGCGCTGAAGTC 1796
Db 1541 GCTTGAAGAGCAGAGGAGGCGCAGAGTGTGACAGGCGGCTCAGCGCTGCGCTGAAGTC 1600

QY 1797 CCGCCAGTCAAGTGTACCCCTCCAGGCGAGTGCCTGTGTGCCCAACCGGACCGGCTT 1856
Db 1601 CCGCCAGTCAAGTGTACCCCTCCAGGCGAGTGCCTGTGTGCCCAACCGGACCGGCTT 1660
QY 1857 CCAGAGCAGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1916
Db 1661 CCAGAGCAGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1720
QY 1917 CAGCTTCCGGAAGTGTGTGAGACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1976
Db 1721 CAGCTTCCGGAAGTGTGTGAGACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1780
QY 1977 CTGA 1980
Db 1781 CTGA 1784

RESULT 12

BC028922

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

REMARK

COMMENT

FEATURES

source

CDS

/codon_start=1

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 30 Row: g Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

Location/Qualifiers

1. 2736

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="MGC:25738 IMAGE:3987714"

/tissue_type="mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."

/clone_lib="NCI CGAP_Lu29"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

93. 2042

/codon_start=1


```

/product="Similar to slingshot 3"
/protein_id="AAH28922.1"
/db_xref="GI:20810383"
/translation="MALVTVSRSPASGHSPTVGPTQDRVVRRRGRLQRQSFVILRG
AVLGLQDGGSNVASEADSEPMEEPSGEEPTEDQTDKGQLQSPWKQYQKRHLHLMV
ELLRPQDDIRLAQLLEARPPRLRYLVYSTGEELSEEAILLGVDFFDSSSHSCTIGL
VLPWSDTQYVLLDGDGFSVTSQGQSRIFFKPSVIQTMMATLQVLHQACEVALGSLVP
GGSALAMATYYQEQELKNSDQGCLENMAMSDLESFRPPNAEPGQASEQEMEQAILAEL
WQVLDTSDLDVTSKEIRQALELRGLCPLOQYRDFIDNQMLLMAQODRASRIFFPLY
LGSEWMNANILELQKNRVSHILNMAREIDNFFPERFTYNNVRVWDEESAQLLPHMKET
HRFIEDARAQGTVLVHCQKSHVSRSAATVLAAYAMKQGWMDLEQALIHQELRPIVRPN
HGFLRQLRTTYQGITLTASRQSHVSEQKYGVSPEEPLAEVSTPLPLPDEPBGSGEVM
VMGLEGSQETPEKEELGLRPRINLRGVMSISLLEPSESESTPEAGGLPEVFSDEEPL
HPFSQLSRAKGGQVRVKGPWPALKSRQSVVALHSALVASTRAFDQEGCGQGEQSEBP
MSSTPLRLRWKVAQASVDSDSREEDKA"

```

Query Match	65.3%	Score 1292.6	DB 10	Length 2736
Best Local Similarity	80.6%	Pred. No. 1.5e-233		
Matches 1598	Conservative 0	Mismatches 349	Indels 36	Gaps 6
QY 1	ATGCGCTGCTCAAGTGAAGCGCTTGGCCCGGCGCAGCGCGCTCCACGCGCTGGGG	60		
Db 93	ATGCGCTGCTCAAGTGAAGCGCTTGGCCCGGCGCAGCGCGCTCCACGCGCTGGGA	152		
QY 61	CCCTG---GACCAGCGCGCTCCAGGAAGAGTGCATCCAGGAAGCAGAGCTTTGCG	117		
Db 153	CCACGACGAGACCGAGTGTGAAGACGACAGAGCGCGCTCCAGCGACGAGTTTGGG	212		
QY 118	GTGTCCTGTGGGGCTGTCTCTGGACTGCAGAGTGAAGGGAACAATGATGATCAGCAGAG	177		
Db 213	GTGTCCTGAGGGGCTGTCTCTGGACTTCAGAGTGAAGGGAACAATGATGATCAGAG	272		
QY 178	GCCAGTTCTGAGCCCAACAGAGAAGCGCCGAGTGAAGAGAGCTCCACGGGACACAGACA	237		
Db 273	GCTGACTCTGAGCCCAATGAGAGAACCTCAGGTGAGAGAACAGCCACCGAGACCAAGACT	332		
QY 238	GACTTCGGGCAAGGATCCACAGATCCCGAGAGCAGAGAGAGAGAGGACAGCTGCAC	297		
Db 333	GACAAAGGGCAAGGACTCCAGAGTCCCTGGAAC--AAGTGCAAAAGCGGACCTGCAC	389		
QY 298	CTCATGTTACAGCTGTGAGGCGCAGAGTGCATCCGCTGCGACGCCAGCTGGAAGCA	357		
Db 390	CTCATGTGTGAGTACTAGAGCGCAGAGATGATATCCGCTGCGACGCCAGCTGGAAGCA	449		
QY 358	CCCCGGCTCCCCGGCTCCGCTACCTGCTGCTGATTTCTACACGAGAGAGAAAGTCTG	417		
Db 450	GCCCCACCCCCCAGACTCCGCTACCTGCTGCTGATGATCCACGGGAGAGAACTGAGT---	506		
QY 418	AGCCAGATGAGACGGTCTCTCTGGGGCGTGATTTCCCTGACACAGCAGCTCCCCAGCTGC	477		
Db 507	-----GAGGAGGCCATCTCTGTTGGGGTGGACTTCCCTGACACAGAGTTCCACAGCTGC	560		
QY 478	ACCTTGAGCTGTCTTGCCCCCTCTGGAGTGCACACCAAGGTGTACTTAGATGAGACGGG	537		
Db 561	ACCTTGAGCTGTCTTGCCCCCTCTGGAGTGCATCCAGGTGTACTTGAGCGCGATGGG	620		
QY 538	GGCTTCAGCGTGACGTCTGTGGGCAAAAGCCGGATCTTCAAGCCCATCTCCATCCAGACC	597		
Db 621	GGCTTCAGTGATCATCTGTGGTCAAAAGCCGAATCTTCAAGCCTGTCTTATCCAGACC	680		
QY 598	ATGTGGCCACACTCCAGGATTTGCAACCAAGCATGTGAGGCGAGCTCTAGGACGGGCTT	657		
Db 681	ATGTGGCCACACTGCAGGATTTGCAACCAAGCATGTGAGGCTAGGAGTGGGCTT	740		
QY 658	GTACCGGGTGGCAGTGCCCTCACTGGGCGCAGCCACTACAGAGAGAGACTGAACTCCGAA	717		
Db 741	GTGCTGTGTGGCAGTGCCCTTGCTGGGCTACTTACTACAGAGAGAACTCAACTCTGAC	800		
QY 718	CAGAGCTGCTCAATGAGTGAAGGCTATGCGCGACCTGGAGTCTCTGCGGCTCCAGC	777		
Db 801	CAGGTTGCTCAATGAGTGAAGGCTATGCTGACCTGGAGTCTTTCGACCACTAAT	860		

QY	778	GCCGAGCCTGGCGGCTCTCTCAGAAACAGGAGCAGATGGAGCAGGCGATCCGCTGCTGAGCTG	837
Db	861	GCTGAGCCTGGACAGGCTTTCAGAGCAGGAAAAAGATGAGCAGGCGATCCTGGCTGAAATTG	920
QY	838	TGGAAAGTGTGGATGTCAGTGAACCTGGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCT	897
Db	921	TGGCAGGTGTTGACACCAAGTGAACCTGGACAGCGTCACTTCTAAAGAGATCCGTCAGGCC	980
QY	898	CTGAGCTGCGCCTGGGGCTCCCCCTCCAGCAGTAACCTGACTTCATGCAACAAGATG	957
Db	981	CTGAGCTGCGTCTGGAGTGTCTCTCCAGCAGTAACCTGACTTATTGACAACAGATG	1040
QY	958	CTGCTGCTGTGGCAGACGGGACCAGGACCTCCCGCATCTTCCCCACCTTACCTGGGC	1017
Db	1041	TTCGTCTCATGCGCCAGCAAGACCGGGCTCTCCGATCTTCCCCACCTTACTTTGGGC	1100
QY	1018	TCAGAGTGAACCGACGAAACCTGGAGAGCTGCAGAGGAACAGGGTACCCACATCTTG	1077
Db	1101	TCTGAGTGAATGCTGCCAACCTGGAGGAATTCAGAAAAACAGAGTAAGTCACTTCCTG	1160
QY	1078	AACATGCCCCGGAGATTGACAATCTTACCCTGACGCTTCACTTACCAAAATGCGC	1137
Db	1161	AACATGCCCCGGAGATTGACAATCTTCCCTGACGCTTCACTTATTAACAATGACGT	1220
QY	1138	CTCTGGGATGAGGAGTGGCCCCAGCTGTGCGGCACTGGAAGAGACGACCGCTTCATT	1197
Db	1221	GTCTGGGATGAGGAATCGGCACAGCTTCTGCCCACTGGAAGAGACACATGATTCATT	1280
QY	1198	GAGGCTGCAAGAGCACAGGGCACCCACGCTGTGTCCTGCAAGATGAGCGTCAAGCCG	1257
Db	1281	GAGGATGCCAGAGCACAGGGCACTCGGGGTGCTAGTCCACTGTAAGATGGGTGTCAGCCGT	1340
QY	1258	TCAGCGGCCACAGTGTGCTGCTATGCCATGAAGCAGTAAGATGACGCTGAGAGAGGCC	1317
Db	1341	TCTGCCGCCACGCTAGCTAGCCTATGCCATGAACAGATAGCTGGAGCCTGAGAGAGCC	1400
QY	1318	CTGCGCCACGTGCAGAGACTCCGGCCCATCGCCCGCCCAACCCTGGCTTCTGCGCCAG	1377
Db	1401	CTGATCCACGTGCAGAGACTCCGGCCCATTTGTAAGCCCCCAACCAAGGCTTCTGCGCCAG	1460
QY	1378	CTGCAGATCTACCAAGGSCATCTGACGGGCCAGCGGCCAGAGCCATGTCTGGAGCAGAA	1437
Db	1461	CTACGGACCTACCAAGGSCATCTGACTGCCAGCCGGCAGAGCCACGCTGGGAGCAGAAA	1520
QY	1438	GTGGGTGGGCTCTCCCCAGAGAGGACACCAGCCCTGAAGTCTTACACCATTTCCCACT	1497
Db	1521	GTGGGTGTGCTCTCCCCAGAGAGGACCCCTGGCGCTGAAGTTTCTACACCATTTGCCACT	1580
QY	1498	CTTCCGCCAGAACCTGAGGGTGTGGGGAGGAGAAAGTGTAGGCATGGAAGAGAGCCAG	1557
Db	1581	CTTCCACCAGAACCCAGGGGCGAGTGGGAGGTGATGTATATGGGCTGGAAGGAAAGCCAG	1640
QY	1558	GCAGCCCCGAAAGAAAGAGCCTTGGGCCACGCGCCACGTAATAAACCCTCCGAGGGGTATGAGG	1617
Db	1641	GAAACCCCAAAGAAAGAACTTGGACTGCGGCCCGCATCAACCTTCGAGGGGTATGCGC	1700
QY	1618	TCCATCAGTCTTCTGAGCCCTCTTGGAGCTGAGAGACCACTCAGAGAGCCAGTGAATG	1677
Db	1701	TCCATCAGTCTCTGAGCCCTC--AGAGTCAAGAGATACCCCAAGAGCGGGAGGCGCTG	1757
QY	1678	CCAGAGGTCTTCTTCCCAAGAGTCTTACATGAAGAGCCTCTGCAGCCCTTCCCAAG	1737
Db	1758	CCAGAGTTTTC-----TCTTCAAGATGAAGAGCCTCTTCACTCCATTTCTCTCAG	1805
QY	1738	CTTGCAAGGACCAAGGAGGCCAGCAGGTGACAGGGGGCCTCAGCCTGCGCTGAAGTCC	1797
Db	1806	CTTTCAGAGCCCAAGGTGGCCAGCGGGTCCGCAAAAGGGCCTTGGCCTGCGCTGAAGTCT	1865
QY	1798	CGCCAGTCAAGTGTAACTCTCCAGAGGCGAGTGCCTGTGGTGGCCAAACCGAGCCAGGCTTC	1857
Db	1866	CGACAGTCTGTGTGCTCTCATAGTGTGCTGCGCTGTGGCCAGCAGAGACAGCGGCTTC	1925

QY 1858 CAGAGCAGAGCAGGGGCGAGGGGCGAGGAGACCCCTGCATTCTCTACGCC 1917
DB 1926 -----CAAGAACAGGGGCAAGGGCAAGACAGAGTGCCTGTATGTCTTACCCCT 1979
QY 1918 AGGTTCCGGAAGGTGTGAGACAGGCCACCGTGCATGACGTGAGAGAGAGGGCGAGGCC 1977
DB 1980 AGGCTCCGGAAGGTGTGAGAGCGGCCAGTGTAGATGACAGCAGGAGAGAGCAAGGCC 2039
QY 1978 TGA 1980
DB 2040 TAA 2042

RESULT 13
LOCUS BC004176 1905 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, similar to hypothetical protein FLJ10928, clone
MGC:2772 IMAGE:2958967, mRNA, complete cds.

ACCESSION BC004176
VERSION BC004176
KEYWORDS BC004176.1 GI:13278815
SOURCE MGC.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1905)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettelman and Anuradha Madan

FEATURES
source
1. 1905
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 2 Row: m Column: 8.
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:2772 IMAGE:2958967"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC_14"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
272..1171
/codon_start=1
/product="Similar to hypothetical protein FLJ10928"
/protein_id="AAH04176.1"
/db_xref="GI:13278816"

CDS
BASE COUNT 418 a 627 c 517 g 343 t
ORIGIN
STPRKRVVRSQSVHDSGEEGA"
APKEPGRPRINLRGVMSISLPSLESTSETSDMEVFSSHSSHEPLQPF
OLARTKGQOVDRGPALRSQSVTLQGSAAVANRTQAFOEQEQGQGGEPCTIS
YOGILTASRSHVMEOKVGSVSEHPAEPSTPPLPPEPEGGEKVVMEESCA
QGTHVIVHCKMVSRSATVLAAMKQYECLEALRHVQLRPIARPNGFLRLQI

Query Match 58.3%; Score 1153.4; DB 9; Length 1905;
Best Local Similarity 99.8%; Pred. No. 2.7e-207;

Matches 1165; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 814 GAGCAGCGCATCCGTGCTGAGCTGTGAAAGTGTGATGTACGTACCTGAGAGTGTG 873
DB 6 GAGAGGCGCATCCGTGCT-AGCTGTGAAAGTGTGATGTACGTACCTGAGAGTGTG 64
QY 874 ACTTCCAAGAGATCCGCCAGGCTGTGAGCTGCGCTGGGGCTCCCCCTCAGCAGTAC 933
DB 65 ACTTCCAAGAGATCCGCCAGGCTGTGAGCTGCGCTGGGGCTCCCCCTCAGCAGTAC 124
QY 934 CGTGAATTCATCGACCAACCAAGATGCTGCTGTGTGACAGCGGAGAGCTCCCGC 993
DB 125 CGTGAATTCATCGACCAACCAAGATGCTGCTGTGTGACAGCGGAGAGCTCCCGC 184
QY 994 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAACCGCAACCACTGAGAGCTGAC 1053
DB 185 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAACCGCAACCACTGAGAGCTGAC 244
QY 1054 AGGACAGAGGTCAACCCACATCTTGAACATGGCCCGGAGATTGACACTTCTACCTGAG 1113
DB 245 AGGACAGAGGTCAACCCACATCTTGAACATGGCCCGGAGATTGACACTTCTACCTGAG 304
QY 1114 CGCTTCACTTACCAACATGTGCGCTCTGGGATGAGAGTCCGCCAGCTGCTCCGCAC 1173
DB 305 CGCTTCACTTACCAACATGTGCGCTCTGGGATGAGAGTCCGCCAGCTGCTCCGCAC 364
QY 1174 TGGAGAGAGAGCCACCGCTTCAATTGAGGCTGCAAGAGACAGGGCACCCACGCTGTC 1233
DB 365 TGGAGAGAGAGCCACCGCTTCAATTGAGGCTGCAAGAGACAGGGCACCCACGCTGTC 424
QY 1234 CACTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGCTATGCCATGAAGCAG 1293
DB 425 CACTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTGTGCTATGCCATGAAGCAG 484
QY 1294 TACGAATGACAGCTGAGCAGAGCCCTGCGCCACGTCAGAGAGCTCCGCCATGCGCCG 1353
DB 485 TACGAATGACAGCTGAGCAGAGCCCTGCGCCACGTCAGAGAGCTCCGCCATGCGCCG 544
QY 1354 CCCAACCTGCTTCTCTGCGCCAGCTGCAAGTCTTACAGGGCATCTGACGGCCAGCCG 1413
DB 545 CCCAACCTGCTTCTCTGCGCCAGCTGCAAGTCTTACAGGGCATCTGACGGCCAGCCG 604
QY 1414 CAGAGCATGTCTGGAGCAGAAAGTGGTGGGTCTCCCAAGAGAGCAGCCAGCCCT 1473
DB 605 CAGAGCATGTCTGGAGCAGAAAGTGGTGGGTCTCCCAAGAGAGCAGCCAGCCCT 664
QY 1474 GAAGTCTTACACCATTCACCACTCTCCGCCAGAACTGAGGCTGTGGGAGAGAGAAG 1533
DB 665 GAAGTCTTACACCATTCACCACTCTCCGCCAGAACTGAGGCTGTGGGAGAGAGAAG 724
QY 1534 GTTGTAGGATGGAAGAGAGCCAGGAGCCCGGAAAGAGAGCTGGGCCAGGCCACGT 1593
DB 725 GTTGTAGGATGGAAGAGAGCCAGGAGCCCGGAAAGAGAGCTGGGCCAGGCCACGT 784
QY 1594 ATAAACCTCCGAGGGGTATGAGGTTCATCACTTCTTGAAGCCCTCTTGAAGCTGAG 1653
DB 785 ATAAACCTCCGAGGGGTATGAGGTTCATCACTTCTTGAAGCCCTCTTGAAGCTGAG 844
QY 1654 AGCACCCTCAGAGACAGTGCATGACATGCCAGAGGTTCTTCTTCCACAGAGTCTTCAATGAA 1713
DB 845 AGCACCCTCAGAGACAGTGCATGACATGCCAGAGGTTCTTCTTCCACAGAGTCTTCAATGAA 904
QY 1714 GAGCCTCTGACAGCCCTTCCACAGCTTTCAGAGACCAAGGAGGCCAGCAGGTGACAGG 1773
DB 905 GAGCCTCTGACAGCCCTTCCACAGCTTTCAGAGACCAAGGAGGCCAGCAGGTGACAGG 964
QY 1774 GGGCCTCAGCCTGCGCTGAAGTCCCGGACAGTGTGTTACCTTCCAGGGCAGTGCCTG 1833
DB 965 GGGCCTCAGCCTGCGCTGAAGTCCCGGACAGTGTGTTACCTTCCAGGGCAGTGCCTG 1024
QY 1834 GTGGCCAACCGAGCCAGGCTTCCAGAGCAGAGCAGGGGCAAGGGGCAAGGGGCAAGGGA 1893
DB 1025 GTGGCCAACCGAGCCAGGCTTCCAGAGCAGAGCAGGGGCAAGGGGCAAGGGGCAAGGGA 1084

QY	1894	GAGCCCTGCATTCTCTCTACGCCCCAGGTTCCCGAAGGTGTGAGACAGGCCAGCGTGCAT	1953
Db	1085	GAGCCCTGCATTCTCTCTACGCCCCAGGTTCCCGAAGGTGTGAGACAGGCCAGCGTGCAT	1144
QY	1954	GACAGTGAGAGAGAGCGGCGAGGCTGA	1980
Db	1145	GACAGTGAGAGAGAGCGGCGAGGCTGA	1171
RESULT 14	BC004210	1905 bp mRNA linear PRI 12-JUL-2001	
LOCUS	BC004210	Homo sapiens, Similar to hypothetical protein FLJ10928, clone	
DEFINITION	MGC:4436 IMAGE:2958967, mRNA, complete cds.		
ACCESSION	BC004210		
VERSION	BC004210.1	GI:13278902	
KEYWORDS	MGC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	1 (bases 1 to 1905)		
TITLE	Strausberg, R.		
JOURNAL	Direct Submission		
REMARK	Submitted (01-MAR-2001) National Institutes of Health, Mammalian		
COMMENT	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
	NIH-MGC Project URL: http://mgc.nci.nih.gov		
	Contact: MGC help desk		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: DCTD/DTP		
	CDNA Library Preparation: Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)		
	DNA Sequencing by: Institute for Systems Biology		
	http://www.systemsbio.org		
	contact: amadan@systemsbiology.org		
	Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia		
	Greene, Mark Ketteman and Anuradha Madan		
FEATURES	Clone distribution: MGC clone distribution information can be found		
source	through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov		
	Series: IRAL Plate: 10 Row: g Column: 1.		
	Location/Qualifiers		
	1. .1905		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="MGC:4436 IMAGE:2958967"		
	/tissue_type="Kidney, renal cell adenocarcinoma"		
	/clone_lib="NIH MGC_14"		
	/lab_host="DH10B-R"		
	/note="Vector: pOTB7"		
	272. .1171		
	/codon_start=1		
	/product="Similar to hypothetical protein FLJ10928"		
	/protein_id="AA04210.1"		
	/db_xref="GI:13278903"		
	/translation="MAREIDNFFPERFTHNVRLWDESAQLPHWKETHRFIEARA		
	QGTHTLVHCKMGVSRSAATVLAAMKQYECLEQALRHVQELRPIARPNGFLRQLOI		
	YQGITASRQSHWEQKVGVSPEHPADEVTPFPLPPEEGGGEKXVGMESQA		
	APKEEPGPRPRLNKGVMRSISLLEPSLEBSTSETSDMPEVFSHSHHEPPLQFP		
	OLARTKGQGVDRGPQALPKRSQSVTLQGSVVANRTOAFQEQEQGQGGQGPICIS		
	STPRFRKVRQASVHDSGEEGEA"		
BASE COUNT	418 a 627 c 517 g 343 t		
ORIGIN			
Query Match	58.3%;	Score 1153.4;	DB 9; Length 1905;
Best Local Similarity	99.8%;	Pred. No. 2.7e-207;	
Matches 1165; Conservative	0;	Mismatches 1;	Indels 1; Gaps 1;
QY	814	GAGCAGGGCATCCGTGCTGAGCTGTGGAAGTGTGGATGTCAGTGACCTGAGAGTGTTC	873

Db	6	GAGAGGCGGATCCGCTGCT-AGCTGTGGAAGTGTGGATGTCAAGTGACCTGAGAGTGTG	64
QY	874	ACTTCCAAAGAGATCCGCCAGGCTCTGGAGCTGCGCTGGGGCTCCCTCCAGCAGTAC	933
Db	65	ACTTCCAAAGAGATCCGCCAGGCTCTGGAGCTGCGCTGGGGCTCCCTCCAGCAGTAC	124
QY	934	CGTACTTCATCGACAAACCAGATGCTGCTGCTGTGGCACAGCGGGAACCGAGCTCCCGC	993
Db	125	CGTACTTCATCGACAAACCAGATGCTGCTGCTGTGGCACAGCGGGAACCGAGCTCCCGC	184
QY	994	ATCTTCCCCCACCCTCTACCTTGGGCTCAGAGTGAACCGACCAACCTGGAGAGCTGCAG	1053
Db	185	ATCTTCCCCCACCCTCTACCTTGGGCTCAGAGTGAACCGACCAACCTGGAGAGCTGCAG	244
QY	1054	AGGAACAGGGTCAACCCACATCTTGAACATGGCCCCGGAGATTGACAACTTTACCTTGAG	1113
Db	245	AGGAACAGGGTCAACCCACATCTTGAACATGGCCCCGGAGATTGACAACTTTACCTTGAG	304
QY	1114	CGCTTCACTTACCACAATGTGCGCTCTGGGATGAGGAGTGGGCCAGCTGCTGCGGCAC	1173
Db	305	CGCTTCACTTACCACAATGTGCGCTCTGGGATGAGGAGTGGGCCAGCTGCTGCGGCAC	364
QY	1174	TGGAAGGAGACGACACCGCTTCATTGAGGCTGCAAGACACAGGGGCAACCACTGTGTGTC	1233
Db	365	TGGAAGGAGACGACACCGCTTCATTGAGGCTGCAAGACACAGGGGCAACCACTGTGTGTC	424
QY	1234	CACTGCAAGATGGGCGCTCAGCGCTCAGCGGCCACAGTGTGCTATGCCATGAAGCAG	1293
Db	425	CACTGCAAGATGGGCGCTCAGCGCTCAGCGGCCACAGTGTGCTATGCCATGAAGCAG	484
QY	1294	TACGAATGCAGCCTGGAGCAGGCCCCCTGCGCCACCTGCAAGAGCTCCGGCCCATCGCCGC	1353
Db	485	TACGAATGCAGCCTGGAGCAGGCCCCCTGCGCCACCTGCAAGAGCTCCGGCCCATCGCCGC	544
QY	1354	CCCAACCCCTGGCTTCTCTGCGCCAGCTGCAGATCTACCAAGGCATCCTGACCGGCCAGCCGC	1413
Db	545	CCCAACCCCTGGCTTCTCTGCGCCAGCTGCAGATCTACCAAGGCATCCTGACCGGCCAGCCGC	604
QY	1414	CAGAGCCATGTCTGGAGCAGAAAGTGGTGGGCTCTCCCAAGAGGACACCCAGCCCTT	1473
Db	605	CAGAGCCATGTCTGGAGCAGAAAGTGGTGGGCTCTCCCAAGAGGACACCCAGCCCTT	664
QY	1474	GAACTCTCTACACCATTCACCCTCTTCCGCCAAGAACCTGAGGGTGTGGGGAGAGAAAG	1533
Db	665	GAACTCTCTACACCATTCACCCTCTTCCGCCAAGAACCTGAGGGTGTGGGGAGAGAAAG	724
QY	1534	GTTGTAGGCATGGAAGAGACCAGGCGACCCCCGAAAGAGAGCCTGGGCCACGGCCACGT	1593
Db	725	GTTGTAGGCATGGAAGAGACCAGGCGACCCCCGAAAGAGAGCCTGGGCCACGGCCACGT	784
QY	1594	ATAAACCTCCGAGGGGTATGAGGTCATCATGCTCTTCTGAGCCCTCCTTGGAGCTGGAG	1653
Db	785	ATAAACCTCCGAGGGGTATGAGGTCATCATGCTCTTCTGAGCCCTCCTTGGAGCTGGAG	844
QY	1654	AGCACTCAGAGACCAGTGACATGCCAGAGGCTTTCTTCTTCCACGAGTCTTCACATGAA	1713
Db	845	AGCACTCAGAGACCAGTGACATGCCAGAGGCTTTCTTCTTCCACGAGTCTTCACATGAA	904
QY	1714	GAGCCTCTGCAGCCCTTCCACAGCTTGCAGAGACCAAGGAGGCCAGCAGGTGACAGG	1773
Db	905	GAGCCTCTGCAGCCCTTCCACAGCTTGCAGAGACCAAGGAGGCCAGCAGGTGACAGG	964
QY	1774	GGGCTCAGCCTGCGCTGGAAGTCCCGCCAGTCACTGTTACCTCCAGGGCAGTGCCTG	1833
Db	965	GGGCTCAGCCTGCGCTGGAAGTCCCGCCAGTCACTGTTACCTCCAGGGCAGTGCCTG	1024
QY	1834	GTTGCCAAACCGAGCCAGGCTTCCAGAGCAGAGACAGGGGCAAGGGGCAAGGGGCAAGGGA	1893
Db	1025	GTTGCCAAACCGAGCCAGGCTTCCAGAGCAGAGACAGGGGCAAGGGGCAAGGGGCAAGGGA	1084
QY	1894	GAGCCTGCATTTCTCTACGCCCCAGGTTCCGGAAGTGTGTAGACAGGCCAGCGTGCAT	1953

Db 1085 GAGCCCTGCATTTCTCTACGCCCCAGGTTCGGGAAGGTGTGAACAGAGCCAGCTTCAT 1144
QY 1954 GACAGTGGAGAGAGGGCGAGGCTGA 1980
Db 1145 GACAGTGGAGAGAGGGCGAGGCTGA 1171

RESULT 15
AX086034 1026 bp DNA linear PAT 09-MAR-2001
LOCUS Sequence 31 from Patent WO0112819.
DEFINITION AX086034
ACCESSION AX086034
VERSION AX086034.1 GI:13275848
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ploewman, G.D., Martinez, R., Whyte, D., Hill, R., Flanagan, P. and
Lioubin, M.
TITLE Protein phosphatases and diagnosis and treatment of
phosphatase-related disorders
JOURNAL Patent: WO 0112819-A 31 22-FEB-2001;
Susgen, Inc. (US)
FEATURES
Location/Qualifiers
1..1026
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 217 a 324 c 326 g 159 t
ORIGIN

Query Match 51.8%; Score 1026; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 3e-183;
Matches 1026; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 955 ATGCTGCTGTGGTGACACAGCGGGAGCGGCTCCCGCATCTTCCCCACCTTCACTG 1014
Db 1 ATGCTGCTGTGGTGACACAGCGGGAGCGGCTCCCGCATCTTCCCCACCTTCACTG 60

QY 1015 GGCTCAGAGTGAACGACGCAAACTTGAGAGAGCTGCAGAGAAACAGGGTACCCACATC 1074
Db 61 GGCTCAGAGTGAACGACGCAAACTTGAGAGAGCTGCAGAGAAACAGGGTACCCACATC 120

QY 1075 TTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATGTG 1134
Db 121 TTGAACATGGCCCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATGTG 180

QY 1135 CGCCTCTGGATGAGAGTGGCCCACTGCTGCGCACTGGAAGAGAGACCGCTTC 1194
Db 181 CGCCTCTGGATGAGAGTGGCCCACTGCTGCGCACTGGAAGAGAGACCGCTTC 240

QY 1195 ATTGAGCTGCAAGACAGACAGGACACGACGTCGTGCTGCACTGCAAGATGGGGTCAGC 1254
Db 241 ATTGAGCTGCAAGACAGACAGGACACGACGTCGTGCTGCACTGCAAGATGGGGTCAGC 300

QY 1255 CGCTCAGCGGCAAGTGTGCTGCTATGCCATGAAGCAGTACGAATGCAGCTTGAAGCAG 1314
Db 301 CGCTCAGCGGCAAGTGTGCTGCTATGCCATGAAGCAGTACGAATGCAGCTTGAAGCAG 360

QY 1315 GCCCTGCGCCACGTGACAGAGCTCCGGCCCATGCCCCCAACCTGGCTTCTGCGC 1374
Db 361 GCCCTGCGCCACGTGACAGAGCTCCGGCCCATGCCCCCAACCTGGCTTCTGCGC 420

QY 1375 CAGCTCAGATCTACCAAGGACATCTGACGGCCAGCCGCAAGCCATGTCTGGAGACAG 1434
Db 421 CAGCTCAGATCTACCAAGGACATCTGACGGCCAGCCGCAAGCCATGTCTGGAGACAG 480

QY 1435 AAAGTGGTGGGTCTCCCAAGAGAGACCCAGCCCTGAAGTCTTACACCATTTCCA 1494
Db 481 AAAGTGGTGGGTCTCCCAAGAGAGACCCAGCCCTGAAGTCTTACACCATTTCCA 540

QY 1495 CTTCTCCGCCAGAACCTGAGAGGTGTGGGGAGAGAAAGTTGTAGCATGAAGAGAC 1554
Db 541 CTTCTCCGCCAGAACCTGAGAGGTGTGGGGAGAGAAAGTTGTAGCATGAAGAGAC 600

QY 1555 CAGGCAGCCCCGAAGAAGAGCCTGGGCCACGGCCACGTATTAACCTCCGAGGGTCA 1614
Db 601 CAGGCAGCCCCGAAGAAGAGCCTGGGCCACGGCCACGTATTAACCTCCGAGGGTCA 660

QY 1615 AGTCCATCAGTCTTCTGAGAGCCCTCCTTGAGAGCTGAGAGACCTCAGAGACAGTAC 1674
Db 661 AGTCCATCAGTCTTCTGAGAGCCCTCCTTGAGAGCTGAGAGACCTCAGAGACAGTAC 720

QY 1675 ATGCCAGAGTCTTCTCTTCCACAGAGTCTTACATGAAGAGCCTTGACAGCCCTTCCA 1734
Db 721 ATGCCAGAGTCTTCTCTTCCACAGAGTCTTACATGAAGAGCCTTGACAGCCCTTCCA 780

QY 1735 CAGCTTGCAAGAGCAAGGAGGAGGCCAGAGGTGACAGGGGGCTCAGCCTGCTGAAG 1794
Db 781 CAGCTTGCAAGAGCAAGGAGGAGGCCAGAGGTGACAGGGGGCTCAGCCTGCTGAAG 840

QY 1795 TCCCGCCAGTCAAGTGTATACCTTCCAGGAGAGTGCCTGTGTGGCCAAACCGAACCCAGGCC 1854
Db 841 TCCCGCCAGTCAAGTGTATACCTTCCAGGAGAGTGCCTGTGTGGCCAAACCGAACCCAGGCC 900

QY 1855 TTCCAGAGCAGAGACAGGGGAGGGGAGGGGAGGGGAGAGAGCCCTGCATTCTCTACG 1914
Db 901 TTCCAGAGCAGAGACAGGGGAGGGGAGGGGAGGGGAGAGAGCCCTGCATTCTCTACG 960

QY 1915 CCCAGTTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACGTGAGAGAGAGGGCAG 1974
Db 961 CCCAGTTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACGTGAGAGAGAGGGCAG 1020

QY 1975 GCCTGA 1980
Db 1021 GCCTGA 1026

Search completed: January 15, 2004, 17:26:11
Job time : 7076 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 17:26:15 ; Search time 658 Seconds
(without alignments)
10606.431 Million cell updates/sec

Title: US-09-955-732-1
Perfect score: 1980
Sequence: 1 atggccctgtgtcacagtcgag.....gagaggagggcgagcctga 1980

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_NEW_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1980	100.0	1980	10	US-09-955-732-1 Sequence 1, Appli
2	1978.4	99.9	2852	10	US-09-761-640-2 Sequence 2, Appli
3	1804.8	91.2	2905	12	US-10-108-260A-74 Sequence 74, Appli
4	1774	89.6	2704	10	US-09-761-640-1 Sequence 1, Appli
5	1774	89.6	2704	10	US-09-761-640-7 Sequence 7, Appli
6	1684	85.1	2322	13	US-10-363-676-10 Sequence 10, Appli
7	1409	71.2	1416	10	US-09-955-732-20 Sequence 20, Appli
8	1345.6	68.0	2540	10	US-09-761-640-3 Sequence 3, Appli
9	1121.4	56.6	2083	13	US-09-814-353-19925 Sequence 19925, A
10	694	35.1	1755	13	US-10-363-676-4 Sequence 4, Appli
11	491	24.8	599	13	US-10-363-676-8 Sequence 8, Appli
12	397	20.1	409	13	US-10-363-676-7 Sequence 7, Appli
13	362	18.3	2433	12	US-10-108-260A-699 Sequence 699, App
14	356.4	18.0	494	13	US-10-363-676-9 Sequence 9, Appli
15	340.4	17.2	1949	9	US-09-775-925-1 Sequence 1, Appli

16	336	17.0	351	13	US-09-814-353-14318 Sequence 14318, A
17	329	16.6	426	13	US-10-363-676-1 Sequence 1, Appli
18	312	15.8	349	13	US-09-814-353-1576 Sequence 1576, Ap
19	312	15.8	349	13	US-09-814-353-7934 Sequence 7934, Ap
20	308.8	15.6	1711	9	US-09-775-925-5 Sequence 5, Appli
21	300.8	15.2	3488	13	US-10-381-333-19 Sequence 19, Appli
22	279.4	14.1	1771	13	US-10-181-590-18 Sequence 18, Appli
23	204	10.3	1052	9	US-09-775-925-7 Sequence 7, Appli
24	163.4	8.3	969	10	US-09-808-701-12 Sequence 12, Appli
25	163.4	8.3	969	12	US-10-240-145-38 Sequence 38, Appli
26	156.6	7.9	969	15	US-10-233-131-12 Sequence 12, Appli
27	156.6	7.9	1450	15	US-10-103-313-78 Sequence 78, Appli
28	101.8	5.1	531	13	US-10-165-272-3 Sequence 3, Appli
29	101.8	5.1	1390	13	US-10-165-272-1 Sequence 1, Appli
30	98.8	5.0	1357	12	US-10-264-049-910 Sequence 910, App
31	98.8	5.0	1685	14	US-10-002-600-51 Sequence 51, Appli
32	89.2	4.5	2139	11	US-09-865-993-3 Sequence 3, Appli
33	89.2	4.5	2400	12	US-10-373-801-13 Sequence 13, Appli
34	89.2	4.5	2473	12	US-10-240-965-24 Sequence 24, Appli
35	89.2	4.5	2479	11	US-09-865-993-10 Sequence 10, Appli
36	87.6	4.4	2470	10	US-09-964-824A-555 Sequence 555, App
37	87.6	4.4	2470	10	US-09-880-107-3310 Sequence 3310, Ap
38	87.6	4.4	2470	13	US-09-960-706-949 Sequence 949, App
39	85.4	4.3	1209	10	US-09-902-941-1918 Sequence 1918, Ap
40	85.4	4.3	1209	10	US-09-849-626-1918 Sequence 1918, Ap
41	85.4	4.3	1209	13	US-10-113-872-1918 Sequence 1918, Ap
42	85.4	4.3	1209	15	US-10-017-754-1918 Sequence 1918, Ap
43	85.4	4.3	1238	10	US-09-736-457-803 Sequence 803, App
44	85.4	4.3	1238	10	US-09-902-941-803 Sequence 803, App
45	85.4	4.3	1238	10	US-09-849-626-803 Sequence 803, App

ALIGNMENTS

RESULT 1
US-09-955-732-1
Sequence 1, Application US/09955732
Publication No. US20020182203A1
GENERAL INFORMATION:
APPLICANT: Lucbe, Ralf M.
TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.433
CURRENT APPLICATION NUMBER: US/09/955,732
CURRENT FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1980
TYPE: DNA
ORGANISM: Homo sapiens
US-09-955-732-1

Query Match	100.0%;	Score 1980;	DB 10;	Length 1980;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1980;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCCCTGTGTACAGTGAGCCGTCGCCCCGGGCGAGCGCGGCTCCACGCGCGTGGG	60	
DB	1	ATGGCCCTGTGTACAGTGAGCCGTCGCCCCGGGCGAGCGCGGCTCCACGCGCGTGGG	60	
QY	61	CCCTGGAGCAGGCGGTCCAGCGAGAGTGACTCCAGCGGAGCAGAGCTTTGCGGTG	120	
DB	61	CCCTGGAGCAGGCGGTCCAGCGAGAGTGACTCCAGCGGAGCAGAGCTTTGCGGTG	120	
QY	121	CTCCGTGGGGTGTCTGTGGAGTGCAGATGGAGGGGACAATGATGATGACAGAGGCC	180	
DB	121	CTCCGTGGGGTGTCTGTGGAGTGCAGATGGAGGGGACAATGATGATGACAGAGGCC	180	
QY	181	AGTTCTGAGCCAAACAGAGAGGCCCCGAGTGAGGAGGAGCTCCAGCGGGACAGACAGAC	240	
DB	181	AGTTCTGAGCCAAACAGAGAGGCCCCGAGTGAGGAGGAGCTCCAGCGGGACAGACAGAC	240	

QY	241	TTGGGGCAAGGATCCCAAGATCCCAAGACAGAGAGACAGAGGACGACACTTGCACTC	300
Db	241	TTGGGCAAGGATCCAGAGTCCCAAGAGCAGAGAGACAGAGGACGACACTTGCACTC	300
QY	301	ATGTACAGCTGCTGAGGCGCGCAGAGATGACATCCGCTGGCAGCCCACTGGAGGCACCC	360
Db	301	ATGTACAGCTGCTGAGGCGCGCAGAGATGACATCCGCTGGCAGCCCACTGGAGGCACCC	360
QY	361	CGGCTCCCCGGCTCCGCTACTCTGTGATTCTTACACGAGAAAGAGAGGTCTGAGC	420
Db	361	CGGCTCCCCGGCTCCGCTACTCTGTGATTCTTACACGAGAAAGAGAGGTCTGAGC	420
QY	421	CAGGATGAGACGGTCTCTCGGCGTGGAATTTCCCTGACAGCAGCTCCCCAGCTGCACC	480
Db	421	CAGGATGAGACGGTCTCTCGGCGTGGAATTTCCCTGACAGCAGCTCCCCAGCTGCACC	480
QY	481	CTGGGCTCGTCTTGCCCCCTCTGAGTGACACCAGGTGTACTTAGATGAGACGGGGGC	540
Db	481	CTGGGCTCGTCTTGCCCCCTCTGAGTGACACCAGGTGTACTTAGATGAGACGGGGGC	540
QY	541	TTACAGCTGACGTCTGTGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG	600
Db	541	TTACAGCTGACGTCTGTGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATG	600
QY	601	TGGGCACTCCAGGTAATGCAACAAGCATGTGAGGCACTCTAGGACAGCGCCTTGTA	660
Db	601	TGGGCACTCCAGGTAATGCAACAAGCATGTGAGGCACTCTAGGACAGCGCCTTGTA	660
QY	661	CCGGGTGAGTGCCCTCACTGGGCGCAGCCACTACAGAGAGACTGAATCCGAACAG	720
Db	661	CCGGGTGAGTGCCCTCACTGGGCGCAGCCACTACAGAGAGACTGAATCCGAACAG	720
QY	721	AGCTGCTCAATGAGTGAGCGGCTATGGCCGACCTGGAGTCTCTGCGGCTCCAGCGCC	780
Db	721	AGCTGCTCAATGAGTGAGCGGCTATGGCCGACCTGGAGTCTCTGCGGCTCCAGCGCC	780
QY	781	GAGCCTGCGGGTCTCAGAAACAGAGCAGATGAGCAGCGCATCCGTCTGAGCTGTGG	840
Db	781	GAGCCTGCGGGTCTCAGAAACAGAGCAGATGAGCAGCGCATCCGTCTGAGCTGTGG	840
QY	841	AAAGTGTGGATGTCACTGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG	900
Db	841	AAAGTGTGGATGTCACTGAGAGTGTCACTTCCAAAGAGATCCGCCAGGCTCTG	900
QY	901	GAGCTGCGCTGGGGTCCCCCTCCAGCAGTACCGTGACTTCATCGACAAACAGATGCTG	960
Db	901	GAGCTGCGCTGGGGTCCCCCTCCAGCAGTACCGTGACTTCATCGACAAACAGATGCTG	960
QY	961	CTGCTGTGGCACAGCGGACCGAGCCTCCCGCATCTTCCCCCACTTACTCTGGGCTCA	1020
Db	961	CTGCTGTGGCACAGCGGACCGAGCCTCCCGCATCTTCCCCCACTTACTCTGGGCTCA	1020
QY	1021	GAGTGAACGACGAAACCTGTGAGGAGCTGCAGAGGAAACAGGGTCACCCACATCTTGAAC	1080
Db	1021	GAGTGAACGACGAAACCTGTGAGGAGCTGCAGAGGAAACAGGGTCACCCACATCTTGAAC	1080
QY	1081	ATGCCCCGGAGATTGACAACTTCAACCTGAGCGCTTCACTTCCCAATATGTGCGCTC	1140
Db	1081	ATGCCCCGGAGATTGACAACTTCAACCTGAGCGCTTCACTTCCCAATATGTGCGCTC	1140
QY	1141	TGGGATGAGGAGTCCGCCACGCTGTCGCCACTGTGAAGGAGACGACCGCTTCATTGAG	1200
Db	1141	TGGGATGAGGAGTCCGCCACGCTGTCGCCACTGTGAAGGAGACGACCGCTTCATTGAG	1200
QY	1201	GCTGCAAGACACAGGGCACCACGTCGTGTTCACTGCAAGATGGGCGTCAAGCCGCTCA	1260
Db	1201	GCTGCAAGACACAGGGCACCACGTCGTGTTCACTGCAAGATGGGCGTCAAGCCGCTCA	1260
QY	1261	GCGGCCACAGTCTGGGCTATGCCATGAAGCAGTACGAATGCAGCTTGAAGCAGGCGCTG	1320
Db	1261	GCGGCCACAGTCTGGGCTATGCCATGAAGCAGTACGAATGCAGCTTGAAGCAGGCGCTG	1320

QY	1321	CGCCACGTGCAGGAGCTCCGGCCCATCGGCCGCCCAACCTTGCTTCCTGCGCAGCTG	1380
Db	1321	CGCCACGTGCAGGAGCTCCGGCCCATCGGCCGCCCAACCTTGCTTCCTGCGCAGCTG	1380
QY	1381	CAGATCTTACCAGGGCATCTCTGACGGCCAGCGCCCAAGCCATGTCTGGGAGCAGAAAGTG	1440
Db	1381	CAGATCTTACCAGGGCATCTCTGACGGCCAGCGCCCAAGCCATGTCTGGGAGCAGAAAGTG	1440
QY	1441	GGTGGGCTCTCCCCAGAGGAGCACCCAGGCCCTGTAAGTCTCTACACCATTTCCCACTCTT	1500
Db	1441	GGTGGGCTCTCCCCAGAGGAGCACCCAGGCCCTGTAAGTCTCTACACCATTTCCCACTCTT	1500
QY	1501	CCGCCAGAACCTGAGGGTGGTGGGAGAGAGAAAGTTGTAGGCATGGAAGAGCCAGGCA	1560
Db	1501	CCGCCAGAACCTGAGGGTGGTGGGAGAGAGAAAGTTGTAGGCATGGAAGAGCCAGGCA	1560
QY	1561	GCCCCGAAAGAAAGAGCCTGGGCCACCGGCCACGTATAAACCTCCGAGGGCTATGAGGTCC	1620
Db	1561	GCCCCGAAAGAAAGAGCCTGGGCCACCGGCCACGTATAAACCTCCGAGGGCTATGAGGTCC	1620
QY	1621	ATCAGTCTTCTGAGGCCCTCTTGAGCTGGAGAGCACCTCAGAGAGCCAGTGAATGCCA	1680
Db	1621	ATCAGTCTTCTGAGGCCCTCTTGAGCTGGAGAGCACCTCAGAGAGCCAGTGAATGCCA	1680
QY	1681	GAGGTCTTCTTCCACGAGTCTTACATGAAGCCTCTGCAGCCCTTCCACAGCTT	1740
Db	1681	GAGGTCTTCTTCCACGAGTCTTACATGAAGCCTCTGCAGCCCTTCCACAGCTT	1740
QY	1741	GCAAGGACCAAGGAGGCCAGCAGGTGACAGGGGGCCTCAGCCTGCCCTGAAGTCCCGC	1800
Db	1741	GCAAGGACCAAGGAGGCCAGCAGGTGACAGGGGGCCTCAGCCTGCCCTGAAGTCCCGC	1800
QY	1801	CAGTCAGTGGTTACCTCCAGGGCAGTGCCTGGTGGCCCAACCGGACCCAGGCCCTTCCAG	1860
Db	1801	CAGTCAGTGGTTACCTCCAGGGCAGTGCCTGGTGGCCCAACCGGACCCAGGCCCTTCCAG	1860
QY	1861	GAGCAGGAGCAGGGGCGAGGGCGAGGGCGAGGGAGAGCCCTGCATTTCTCTTACGCCCAGG	1920
Db	1861	GAGCAGGAGCAGGGGCGAGGGCGAGGGCGAGGGAGAGCCCTGCATTTCTCTTACGCCCAGG	1920
QY	1921	TTCCGGAAGGTGGTGAGACAGGCCAGCGTGCATGACAGTGGAGAGGAGGGCGAGCCCTGA	1980
Db	1921	TTCCGGAAGGTGGTGAGACAGGCCAGCGTGCATGACAGTGGAGAGGAGGGCGAGCCCTGA	1980

```

RESULT 2
US-09-761-640-2
; Sequence 2, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C1000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2852
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-2

```

```

Query Match          99.9%; Score 1978.4; DB 10; Length 2852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy      1  ATGGCCCTGTGCACAGTGAAGCCGTTGCCCGCCGGGACAGCGGCGGCGCTTCCACGCCCGTGGGG 60
      |||
Db      57  ATGGCCCTGTGCACAGTGAAGCCGTTGCCCGCCGGGACAGCGGCGGCGCTTCCACGCCCGTGGGG 116

```

QY 61 CCTGGGACCAGCGGCTCCAGCGAAGAGTGAATCCAGCGAAGCGAGCCTTTGGGCTG 120
| | | | |
Db 117 CCTGGGACCAGCGGCTCCAGCGAAGAGTGAATCCAGCGAAGCGAGCCTTTGGGCTG 176
QY 121 CTCGGTGGGGCTGTCTCGGGAATGACGATGAGGGGACAATGATGATGCAGAGGCC 180
| | | | |
Db 177 CTCGGTGGGGCTGTCTCGGGAATGACGATGAGGGGACAATGATGATGCAGAGGCC 236
QY 181 AGTTCTGAGCCACAGAGAGAGCGCCGAGTGAAGAGAGCTCCAGCGGGAACAGACAGAC 240
| | | | |
Db 237 AGTTCTGAGCCACAGAGAGAGCGCCGAGTGAAGAGAGCTCCAGCGGGAACAGACAGAC 296
QY 241 TTCGGGCAAGGATCCCAAGTCCCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
| | | | |
Db 297 TTCGGGCAAGGATCCCAAGTCCCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 356
QY 301 ATGGTACAGCTGTCTGAGGCGCCAGGATGACATCCGCTTGCGAGCCGAGCTGGAGGCACC 360
| | | | |
Db 357 ATGGTACAGCTGTCTGAGGCGCCAGGATGACATCCGCTTGCGAGCCGAGCTGGAGGCACC 416
QY 361 CGGCTCTCCCGGCTCCGCTACCTGCTGTAGTTTCTACACGAGAGAGAGAGAGAGAGAG 420
| | | | |
Db 417 CGGCTCTCCCGGCTCCGCTACCTGCTGTAGTTTCTACACGAGAGAGAGAGAGAGAGAG 476
QY 421 CAGGATGAGACGGTCTCTCTGCGGCGTGGATTTCCCTGACAGCAGCTCCCCAGCTGCACC 480
| | | | |
Db 477 CAGGATGAGACGGTCTCTCTGCGGCGTGGATTTCCCTGACAGCAGCTCCCCAGCTGCACC 536
QY 481 CTGGGCTCTGCTTGGCCCTCTGAGTGAACACCCAGGTGTACTTAGATGAGAGCGGGGCG 540
| | | | |
Db 537 CTGGGCTCTGCTTGGCCCTCTGAGTGAACACCCAGGTGTACTTAGATGAGAGCGGGGCG 596
QY 541 TTCAGCGTGAAGTCTGTGGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATTG 600
| | | | |
Db 597 TTCAGCGTGAAGTCTGTGGGGCAAAAGCCGATCTTCAAGCCCATCTCCATCCAGACCATTG 656
QY 601 TGGGACCACACTCCAGGTAATGCAACCAAGCATGTAGGCGAGCTCTAGGCAAGCGGCTTGTG 660
| | | | |
Db 657 TGGGACCACACTCCAGGTAATGCAACCAAGCATGTAGGCGAGCTCTAGGCAAGCGGCTTGTG 716
QY 661 CCGGGTGGCAGTGCCTCACTGGGCGCAGCACTACAGAGAGAGAGAGAGAGAGAGAGAGAG 720
| | | | |
Db 717 CCGGGTGGCAGTGCCTCACTGGGCGCAGCACTACAGAGAGAGAGAGAGAGAGAGAGAGAG 776
QY 721 AGCTGCTCAATGATGAGTGAAGCGGCTATAGGCGGCACTGGAGTCTCTGGGCTTCCAGCGCC 780
| | | | |
Db 777 AGCTGCTCAATGATGAGTGAAGCGGCTATAGGCGGCACTGGAGTCTCTGGGCTTCCAGCGCC 836
QY 781 GAGCCTGGCGGCTCTCTAGAAACAGAGACAGATGGAGCAGCGATCCGTGCTGAGCTGTGG 840
| | | | |
Db 837 GAGCCTGGCGGCTCTCTAGAAACAGAGACAGATGGAGCAGCGATCCGTGCTGAGCTGTGG 896
QY 841 AAAGTGTGATGTCAGTGACCTGGAGAGTCACTTCCAAAGAGATCCGCCAGGCTCTG 900
| | | | |
Db 897 AAAGTGTGATGTCAGTGACCTGGAGAGTCACTTCCAAAGAGATCCGCCAGGCTCTG 956
QY 901 GAGCTGGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAACAGAGATGCTG 960
| | | | |
Db 957 GAGCTGGCGCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGACAACAGAGATGCTG 1016
QY 961 CTGCTGTGGCACAAGCGGGAACCGAGCTCCCGCATTTCCGCCACTCTACCTGGGCTCA 1020
| | | | |
Db 1017 CTGCTGTGGCACAAGCGGGAACCGAGCTCCCGCATTTCCGCCACTCTACCTGGGCTCA 1076
QY 1021 GAGTGAACGCAAGCAAACTGGAGGAGCTGCAAGGAACAGGGGTACCCACATCTTGAAC 1080
| | | | |
Db 1077 GAGTGAACGCAAGCAAACTGGAGGAGCTGCAAGGAACAGGGGTACCCACATCTTGAAC 1136
QY 1081 ATGGCCCCGGGAGATTGACAATTCTACCTGAGCGCTTCACTTACCAATGTGCGGCTTC 1140
| | | | |
Db 1137 ATGGCCCCGGGAGATTGACAATTCTACCTGAGCGCTTCACTTACCAATGTGCGGCTTC 1196
QY 1141 TGGGATGAGGAGTGGGCCAGCTGTGCGGCACTGGAAGAGAGAGAGAGAGAGAGAGAGAG 1200

Db 1197 TGGGATGAGGAGTCGGCCCAAGTGTCTGCGGCACTGGAAGAGAGAGAGAGAGAGAGAGAG 1256
QY 1201 GCTGACAGACACAGGGGACACCCACAGTGTGTCTGCTCACTGCAAGATGGGCGCTCAAGCCGCTCA 1260
| | | | |
Db 1257 GCTGACAGACACAGGGGACACCCACAGTGTGTCTGCTGCTGCAAGATGGGCGCTCAAGCCGCTCA 1316
QY 1261 GCGGCCACAGTGTGTGCTTATGCCATGAAGCAGTACGAATGCAGCCTTGAGCAGGCCCTG 1320
| | | | |
Db 1317 GCGGCCACAGTGTGTGCTTATGCCATGAAGCAGTACGAATGCAGCCTTGAGCAGGCCCTG 1376
QY 1321 CGCCACGTGACAGAGCTCCGGGCCATCGCCCCGCCCAACCTTGCTTCTGCGCCAGCTG 1380
| | | | |
Db 1377 CGCCACGTGACAGAGCTCCGGGCCATCGCCCCGCCCAACCTTGCTTCTGCGCCAGCTG 1436
QY 1381 CAGATCTACCAAGGCACTCTGACCGGCCAGCCGCGCAGAGCCATGTCTGGAGCAGAAAGTG 1440
| | | | |
Db 1437 CAGATCTACCAAGGCACTCTGACCGGCCAGCCGCGCAGAGCCATGTCTGGAGCAGAAAGTG 1496
QY 1441 GGTGGGTCTCCCAAGAGAGCAACCCAGCCCTGAAGTCTCTACACCACTTCCACCTCTT 1500
| | | | |
Db 1497 GGTGGGTCTCCCAAGAGAGCAACCCAGCCCTGAAGTCTCTACACCACTTCCACCTCTT 1556
QY 1501 CCGCCAGAACCTGAGGGTGTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
| | | | |
Db 1557 CCGCCAGAACCTGAGGGTGTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1616
QY 1561 GCCCCGAAAGAGAGCGTGGGCCACCGGCCAGTATAAAGCTCCGAGGGGTGATGAGGTCC 1620
| | | | |
Db 1617 GCCCCGAAAGAGAGCGTGGGCCACCGGCCAGTATAAAGCTCCGAGGGGTGATGAGGTCC 1676
QY 1621 ATCAGTCTTCTGAGAGCCCTCTTGAGAGCTGAGAGACACCTCAGAGACCAAGTGAATGCCA 1680
| | | | |
Db 1677 ATCAGTCTTCTGAGAGCCCTCTTGAGAGCTGAGAGACACCTCAGAGACCAAGTGAATGCCA 1736
QY 1681 GAGTCTTCTCTTCCACAGAGTCTTACATGAAGCCCTTGACGCCCTTCCACAGCTT 1740
| | | | |
Db 1737 GAGTCTTCTCTTCCACAGAGTCTTACATGAAGCCCTTGACGCCCTTCCACAGCTT 1796
QY 1741 GCAAGGACCAAGGGAGGGCCACAGGTGACAGGGGCTTCAGCCTGCTGAGTCCCGC 1800
| | | | |
Db 1797 GCAAGGACCAAGGGAGGGCCACAGGTGACAGGGGCTTCAGCCTGCTGAGTCCCGC 1856
QY 1801 CAGTCACTGTACCTTCCAGGGCAGTGCCTGTGTGCCCAACCCGACCCAGGCTTCCAG 1860
| | | | |
Db 1857 CAGTCACTGTACCTTCCAGGGCAGTGCCTGTGTGCCCAACCCGACCCAGGCTTCCAG 1916
QY 1861 GAGCAGAGCAGGGGCGAGGGGCAAGGGGCAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
| | | | |
Db 1917 GAGCAGAGCAGGGGCGAGGGGCAAGGGGCAAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1976
QY 1921 TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGAGAGAGAGAGAGAGAG 1980
| | | | |
Db 1977 TTCCGGAAGGTGTGAGACAGGCCAGCGTGCATGACAGTGAAGAGAGAGAGAGAGAGAGAGAG 2036

RESULT 3
US-10-108-260A-74
; Sequence 74, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 2905
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-74

Query Match 91.2%; Score 1804.8; DB 12; Length 2905;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1909; Conservative 0; Mismatches 2; Indels 83; Gaps 2;

QY 66 GGACCAAGCGGCTCCAGCAAGAGTGCCTCCAGCGAAGGAGAGCTTTGCGTGTCCG 125
DB 198 GGACCAAGCGGCTCCAGCAAGAGTGCCTCCAGCGAAGGAGAGCTTTGCGTGTCCG 253
QY 126 TGGGCTGTCTGGGACTGCAGATGGAAGGAGCAATGATGATGACAGAGGCCAGTTC 185
DB 254 TGGGCTGTCTGGGACTGCAGATGGAAGGAGCAATGATGATGACAGAGGCCAGTTC 313
QY 186 TGAGCCACAGAGAGAGGCGCCGAGTGAAGAGAGAGCTCCAGGGAGACAGACTTCG 245
DB 314 TGAGCCACAGAGAGAGGCGCCGAGTGAAGAGAGAGCTCCAGGGAGACAGACTTCG 373
QY 246 GCAAGGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 305
DB 374 GCAAGGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
QY 306 ACAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 365
DB 434 ACAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 493
QY 366 TCCCGGCTCGGCTACCTGTGTGATTTCTACACGAGAGAGAGAGAGAGAGAGAGAG 425
DB 494 TCCCGGCTCGGCTACCTGTGTGATTTCTACACGAGAGAGAGAGAGAGAGAGAGAG 553
QY 426 TGAGAGCGGTCTCTGGGCGTGGATTTCCCTGACAG-----CAGCT 461
DB 554 TGAGAGCGGTCTCTGGGCGTGGATTTCCCTGACAGAGAGAGAGAGAGAGAGAGAG 613
QY 462 -----CAGCT 466
DB 614 GAGGGCAAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 673
QY 467 CCCCCAGCTGCACTCTGGGCGTGTCTTGCCTCTGAGTGAACACCCAGGTGTAATTAG 526
DB 674 CCCCCAGCTGCACTCTGGGCGTGTCTTGCCTCTGAGTGAACACCCAGGTGTAATTAG 733
QY 527 ATGAGAGCGGCGCTTCAAGCGTGAAGTGTGGGCAAGAGAGAGAGAGAGAGAGAGAG 586
DB 734 ATGAGAGCGGCGCTTCAAGCGTGAAGTGTGGGCAAGAGAGAGAGAGAGAGAGAGAG 793
QY 587 CCATCCAGAGCATGTGGGCAACATCCAGTATGACCAAGCATGTGAGGAGAGAGAGAG 646
DB 794 CCATCCAGAGCATGTGGGCAACATCCAGTATGACCAAGCATGTGAGGAGAGAGAGAG 853
QY 647 GCAGCGGCTGTGATCCGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 706
DB 854 GCAGCGGCTGTGATCCGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
QY 707 TGAAGTCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 766
DB 914 TGAAGTCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973
QY 767 GGGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 826
DB 974 GGGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1033
QY 827 GTGCTGAGCTGTGAAAGTGTGATGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 886
DB 1034 GTGCTGAGCTGTGAAAGTGTGATGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1093
QY 887 TCCGCCAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
DB 1094 TCCGCCAGGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1153
QY 947 ACAACCAAGATGTGCTGTGTGTCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1006
DB 1154 ACAACCAAGATGTGCTGTGTGTCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213

QY 1007 TCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGAGCTGCAGAGAGAGAGAGTCA 1066
DB 1214 TCTACCTGGGCTCAGAGTGAAGCAGCAAACTTGAGAGAGCTGCAGAGAGAGAGAGTCA 1273
QY 1067 CCCACATCTTGAACATGAGCGCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACC 1126
DB 1274 CCCACATCTTGAACATGAGCGCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACC 1333
QY 1127 ACAATGTGCGCTCTGGATGAGAGAGTGGCCAGCTGTGCGGCACTGGAAGAGAGAGAG 1186
DB 1334 ACAATGTGCGCTCTGGATGAGAGAGTGGCCAGCTGTGCGGCACTGGAAGAGAGAGAG 1393
QY 1187 ACCGCTTCAATTGAGGCTGCAAGAGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1246
DB 1394 ACCGCTTCAATTGAGGCTGCAAGAGACAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1453
QY 1247 GCGTCAGCGGCTCAGCGGCGCACAGTGTGGCTATGCAATGAAGAGAGAGAGAGAGAG 1306
DB 1454 GCGTCAGCGGCTCAGCGGCGCACAGTGTGGCTATGCAATGAAGAGAGAGAGAGAGAG 1513
QY 1307 TGGAGCAGGCGCTCGGCGCACAGTGTGGCTATGCAATGAAGAGAGAGAGAGAGAGAG 1366
DB 1514 TGGAGCAGGCGCTCGGCGCACAGTGTGGCTATGCAATGAAGAGAGAGAGAGAGAGAG 1573
QY 1367 TCCTGCGCGAGCTGCAAGATCTACAGGCGCATCTGACGCGCAAGCGCGCAAGCCATGTCT 1426
DB 1574 TCCTGCGCGAGCTGCAAGATCTACAGGCGCATCTGACGCGCAAGCGCGCAAGCCATGTCT 1633
QY 1427 GGGAGCAGAAAGTGGGTGGGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1486
DB 1634 GGGAGCAGAAAGTGGGTGGGTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1693
QY 1487 CATTCACACCTCTTCCGCGCAGAACTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAG 1546
DB 1694 CATTCACACCTCTTCCGCGCAGAACTGAGGGTGTGGGAGAGAGAGAGAGAGAGAGAGAG 1753
QY 1547 AAGAGAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1606
DB 1754 AAGAGAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1813
QY 1607 GGGTCATGAGTCCATCAGTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1666
DB 1814 GGGTCATGAGTCCATCAGTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1873
QY 1667 CCAGTGAATGCGCAGAGAGTCTTCTTCCACAGAGTCTTCAATGAAGAGAGAGAGAGAGAG 1726
DB 1874 CCAGTGAATGCGCAGAGAGTCTTCTTCCACAGAGTCTTCAATGAAGAGAGAGAGAGAGAG 1933
QY 1727 CCTTCCACAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1786
DB 1934 CCTTCCACAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1993
QY 1787 CCCTGAAGTCCCGCAGTCAAGTGTACCTTCCAGGCGAGTGCCTGTGGCCAAACCGGA 1846
DB 1994 CCCTGAAGTCCCGCAGTCAAGTGTACCTTCCAGGCGAGTGCCTGTGGCCAAACCGGA 2053
QY 1847 CCCAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1906
DB 2054 CCCAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2113
QY 1907 CCTCTACGCGCAGGTTCCCGAAGGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1966
DB 2114 CCTCTACGCGCAGGTTCCCGAAGGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2173
QY 1967 AGGGCGAGGCTGA 1980
DB 2174 AGGGCGAGGCTGA 2187

RESULT 4
US-09-761-640-1
; Sequence 1, Application US/09761640
; Patent No. US20020137042A1

GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2704
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-1

Query Match 89.6%; Score 1774; DB 10; Length 2704;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1882; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

QY 1 ATGGCCCTGTACAGTGAAGCCGTTGCCCCCGGGCAGCGCGCTCCACGCCCTGGGG 60
Db 94 ATGGCCCTGTACAGTGAAGCCGTTGCCCCCGGGCAGCGCGCTCCACGCCCTGGGG 153
QY 61 CCTGGGAACAGCGCGTCCAGCGAAGAGTGAAGTCCAGCGAAGCAGAGCTTTGGCGTG 120
Db 154 CCTGGGAACAGCGCGTCCAGCGAAGAGTGAAGTCCAGCGAAGCAGAGCTTTGGCGTG 213
QY 121 CTCCTGGGGCTGTCTGGGAAGTGAAGTGAAGGAGGAGCAATGATGACAGAGGCC 180
Db 214 CTCCTGGGGCTGTCTGGGAAGTGAAGTGAAGGAGGAGCAATGATGACAGAGGCC 273
QY 181 AGTTCTGAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 274 AGTTCTGAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 333
QY 241 TTGGGCAAGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 334 TTGGGCAAGATCCCAAGAGTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 393
QY 301 ATGTACAGCTGTGAAGCCGCAAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 360
Db 394 ATGTACAGCTGTGAAGCCGCAAGATGACATCCGCTGGCAGCCAGCTGAGGCAACC 453
QY 361 CGGCTCCCCGCTCCGCTACCTGTGTGTTTCTACACGAGAGAGAGAGAGAGAGAG 420
Db 454 CGGCTCCCCGCTCCGCTACCTGTGTGTTTCTACACGAGAGAGAGAGAGAGAGAG 513
QY 421 CAGATGAGACGGTCTCTCTGGCGTGGGCTGGATTTCCCTGACAGCAGTCCCCAGCTGCACC 480
Db 514 CAGATGAGACGGTCTCTCTGGCGTGGGCTGGATTTCCCTGACAGCAGTCCCCAGCTGCACC 573
QY 481 CTGGGGCTGTCTTGGCCCTCTGAGTGAACCCAGGTGTACTTAGATGAGACGGGGGC 540
Db 574 CTGGGGCTGTCTTGGCCCTCTGAGTGAACCCAGGTGTACTTAGATGAGACGGGGGC 633
QY 541 TTCAGCGTGAAGTGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATG 600
Db 634 TTCAGCGTGAAGTGTGGGCAAGCGGATCTTCAAGCCCATCTCCATCCAGACCATG 693
QY 601 TGGGCCACACTCCAGTATTTGCAACCAAGCATGTAGGCAAGCTCTAGGCAAGGCTTTGTA 660
Db 694 TGGGCCACACTCCAGTATTTGCAACCAAGCATGTAGGCAAGCTCTAGGCAAGGCTTTGTA 753
QY 661 CCGGCTGAGTGCCTTCACTGGGCAAGCACTACAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 754 CCGGCTGAGTGCCTTCACTGGGCAAGCACTACAGAGAGAGAGAGAGAGAGAGAGAG 813
QY 721 AGTGCCTCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 814 AGTGCCTCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 873
QY 781 GAGCCTGGCGGCTCTCAGAACAGAGAGAGATGAGCAGGCGATCCGTGCTGAGCTGTGG 840

Db 874 GAGCCTGGCGGCTCTCAGAACAGAGAGAGATGAGCAGGCGATCCGTGCTGAGCTGTGG 933
QY 841 AAAGTTGATGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 900
Db 934 AAAGTTGATGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 993
QY 901 GAGCTGGCCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGCAACCAAGATGCTG 960
Db 994 GAGCTGGCCTGGGGCTCCCCCTCCAGCAGTACCGTGAATTCATGCAACCAAGATGCTG 1053
QY 961 CTGCTGTGGCAGACCGGACCGAGCCTCCCGCATCTTCCCCACCTTACCTGGGCTCA 1020
Db 1054 CTGCTGTGGCAGACCGGACCGAGCCTCCCGCATCTTCCCCACCTTACCTGGGCTCA 1113
QY 1021 GAGTGAAACGACGAAACCTGAGAGAGAGTGAAGAGAAACAGGGTCAACCATCTTGAAC 1080
Db 1114 GAGTGAAACGACGAAACCTGAGAGAGAGTGAAGAGAAACAGGGTCAACCATCTTGAAC 1173
QY 1081 ATGGCCGGGAGATTGAACAATTCTACCTGAGCGCTTACCTACCAATGTGCGCCTC 1140
Db 1174 ATGGCCGGGAGATTGAACAATTCTACCTGAGCGCTTACCTACCAATGTGCGCCTC 1233
QY 1141 TGGGATGAGAGTGGGCCAGCTGTGCTGCGCAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1234 TGGGATGAGAGTGGGCCAGCTGTGCTGCGCAGCTGGAAGAGAGAGAGAGAGAGAGAGAG 1293
QY 1201 GCTGCAAGACACAGGGCAACCAAGTGTGTGCTGCTGCAAGATGGGCGTCAAGCCGCTCA 1260
Db 1294 GCTGCAAGACACAGGGCAACCAAGTGTGTGCTGCTGCAAGATGGGCGTCAAGCCGCTCA 1353
QY 1261 GCGGCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1354 GCGGCCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1413
QY 1321 CGCCAGTGAAGAGAGTCCGCGCCCATGCGCCCGCCCAACCTGCTTCTGCGCCAGCTG 1380
Db 1414 CGCCAGTGAAGAGAGTCCGCGCCCATGCGCCCGCCCAACCTGCTTCTGCGCCAGCTG 1473
QY 1381 CAGATTAACAGGGCATCTGACCGGCCAGCCGCAAGCCATGTCTGGGAGCAGAAAGTG 1440
Db 1474 CAGATTAACAGGGCATCTGACG----- 1497
QY 1441 GGTGGGTCTCCCAAGAGAGACCCAGCCCTGAAGTCTTACACCATTTCCCACTCTT 1500
Db 1498 ----- 1497
QY 1501 CCGCCAGAACTGAGGGTGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1498 --GCCAGAACTGAGGGTGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1555
QY 1561 GCGCCGAAAGAGAGAGCTGGGCAAGCGGCAAGTAAACCTCCGAGGGGTGATGAGGTCC 1620
Db 1556 GCGCCGAAAGAGAGAGCTGGGCAAGCGGCAAGTAAACCTCCGAGGGGTGATGAGGTCC 1615
QY 1621 ATCAGTCTTGTGAGACCTCTTGAAGTGAAGCACTCAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1616 ATCAGTCTTGTGAGACCTCTTGAAGTGAAGCACTCAGAGAGAGAGAGAGAGAGAGAG 1675
QY 1681 GAGTCTTCTCTTCCACAGAGTCTTCAATGAAGAGCTCTGACGCCCTTCCACAGCTT 1740
Db 1676 GAGTCTTCTCTTCCACAGAGTCTTCAATGAAGAGCTCTGACGCCCTTCCACAGCTT 1735
QY 1741 GCAAGAACCAAGGAGGCGCAGAGTGAACAGGGGCTCAGCCTGCCCCGAAGTCCCGC 1800
Db 1736 GCAAGAACCAAGGAGGCGCAGAGTGAACAGGGGCTCAGCCTGCCCCGAAGTCCCGC 1795
QY 1801 CAGTCAAGTGTACCTTCCAGGGCAGTGCCTGTGTGGCCAAACCGGACCCAGGCTTCCAG 1860
Db 1796 CAGTCAAGTGTACCTTCCAGGGCAGTGCCTGTGTGGCCAAACCGGACCCAGGCTTCCAG 1855
QY 1861 GAGCAGAGAGAGGGGAGAGGGGAGAGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920

QY 1741 GCAAGGACCAAGGAGGAGCCAGAGTGAGACAGGGGGCCCTCAGCCTGCCCCCTGAAGTCCCCG 1800
Db 1736 GCAAGGACCAAGGAGGAGCCAGAGTGAGACAGGGGGCCCTCAGCCTGCCCCCTGAAGTCCCCG 1795
QY 1801 CAGTCAGTGGTTACCCCTCCAGGGGAGTGCCTGGTGGCCCAACCGGACCCAGGCTTCCAG 1860
Db 1796 CAGTCAGTGGTTACCCCTCCAGGGGAGTGCCTGGTGGCCCAACCGGACCCAGGCTTCCAG 1855
QY 1861 GAGCAGGAGCAGGGGCGAGGGGCGAGGGGAGAGAGCCCTGCATTCTCTACGCCAGG 1920
Db 1856 GAGCAGGAGCAGGGGCGAGGGGCGAGGGGAGAGAGCCCTGCATTCTCTACGCCAGG 1915
QY 1921 TTCCGGAGAGTGTGTGACACAGGCCAGCGTGCATGACAGTGAGAGAGGGCGAGGCTGA 1980
Db 1916 TTCCGGAGAGTGTGTGACACAGGCCAGCGTGCATGACAGTGAGAGAGGGCGAGGCTGA 1975

RESULT 6
US-10-363-676-10
; Sequence 10, Application US/10363676
; Publication No. US20030170856A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN MAP KINASE PHOSPHATASE-LIKE ENZYME
; FILE REFERENCE: L10122 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/363,676
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US 60/230,709
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 2322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-363-676-10

Query Match 85.1%; Score 1684; DB 13; Length 2322;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1852; Conservative 0; Mismatches 75; Indels 23; Gaps 10;
QY 39 CGGCGCTCCACGCGCCCTGGGGCCCTGGAGCCAGGCGGTCCAGCGAGAGTGCATCCA 98
Db 147 CGCCCTCACCCCTGGGGCTGCTCTCTCGCAGAGACTGTCTCAGCGAAGAGTGCATCCA 206
QY 99 GCGAAGCAGAGCTTTGGGGTGTCC-GTGGGCTGTCTCGGACTGCAGATGAGGGG 157
Db 207 GCGA---AGAGCTTTGGGTGTCTCTGTGGGCTGTCTCGGACTGCAGATGAGGGG 262
QY 158 ACAATGATGATGACAGAGAGCCAGTCTTGAGCCACAGAGAGAGGCCCCGAGTGAGAGG 217
Db 263 ACAATGATGATGACAGAGAGCCAGTCTTGAGCCACAGAGAGAGGCCCCGAGTGAGAGG 322
QY 218 AGCTCCACGGG--ACCAAGACAGACTT--CGGGCAAGGATCCCAAGTCCCCAGAGCA 272
Db 323 AGCTCCACGGGGGAACCAAGACAGACTTTCGGTGCAAGGATCCCAAGTCCCCAGAGCA 382
QY 273 GGAAGAGCAGAGGCGAGACCTGCACCTCATGTGTACAGTGTGAGGCCGCAAGATGACAT 332
Db 383 GGAAGAGCAGAGGCGAGCACTGCAACCTCATGTAGGGCTGTGAGGCCGCAAGGATGACAT 442
QY 333 CGGCCTGGCA--GCCCACTGGAGGCAACCCCGGCTCCCGGCTCCGCTACCTGC----- 385
Db 443 CGCACTTGAAGCCCAAGCTGGAAGGCAACCCCGGCTCCCGGATCCGATACCTTGCT 502
QY 386 TGGTAGTTTCTACACGAGAA-GGAGAAAGTGTGAG-CCAGATGAGACGGTCTCTGGG 443
Db 503 TGGTAGTTTCTACACGAGAAAGGAGTGTGAGCCCAAGATGAGACGGTCTCTGGG 562
QY 444 -CGTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACCCCTGGGCTGGTCTTGCCCTCT 502
Db 563 ACGTGATTTCCCTGACAGCAGCTCCCCCAGCTGCACCCCTGGGCTGGTCTTGCCCTCT 622

QY 503 GAGTGACACCCAGGTGTACTTAGATGAGACCGGGGCTTCAGCGTGAAGTCTGTGGGC 562
Db 623 GAGTGACACCCAGGTGTACTTAGATGAGACCGGGGCTTCAGCGTGAAGTCTGTGGGC 682
QY 563 AAAGCCGATCTTCAAGCCATCTCCATCCAGACCATGTGGCCACACTCCAGTATTGC 622
Db 683 AAAGCCGATCTTCAAGCCATCTCCATCCAGACCATGTGGCCACACTCCAGTATTGC 742
QY 623 ACCAAGCATGTGAGGCGAGCTCTAGGCAAGCGGCTTGTACCAGGTGGCTCCACT 682
Db 743 ACCAAGCATGTGAGGCGAGCTCTAGGCAAGCGGCTTGTACCAGGTGGCTCCACT 802
QY 683 GGGCCAGCCACTACCAAGAGAGACTGAACCTCCGAACAGAGCTGCCTCAATGAGTGACGG 742
Db 803 GGGCCAGCCACTACCAAGAGAGACTGAACCTCCGAACAGAGCTGCCTCAATGAGTGACGG 862
QY 743 CTATGCCGACCTGGAGTCTCTGCGGCTCCAGCGCCGAGCCTGGCGGCTCTCAGAAC 802
Db 863 CTATGCCGACCTGGAGTCTCTGCGGCTCCAGCGCCGAGCCTGGCGGCTCTCAGAAC 922
QY 803 AGAGCAGATGAGCAGAGCGCATCCGTGCTGAGCTGTGAAAGTGTGATGTGAGTGACC 862
Db 923 AGAGCAGATGAGCAGAGCGCATCCGTGCTGAGCTGTGAAAGTGTGATGTGAGTGACC 982
QY 863 TGAAGTGTCACTTCCAAAGAGATCCGCAAGGCTCTGAGCTGCCTGGGCTCCCC 922
Db 983 TGAAGTGTCACTTCCAAAGAGATCCGCAAGGCTCTGAGCTGCCTGGGCTCCCC 1042
QY 923 TCCAGAGTACCCTGACTTCATCGACAACCAAGATGCTGCTGTGGGACAGCGGAGC 982
Db 1043 TCCAGAGTACCCTGACTTCATCGACAACCAAGATGCTGCTGTGGGACAGCGGAGC 1102
QY 983 GAGCTTCCCGCATCTTCCCGCACTCTACCTGGGCTCAGAGTGGAACGCAAAACCTGG 1042
Db 1103 GAGCTTCCCGCATCTTCCCGCACTCTACCTGGGCTCAGAGTGGAACGCAAAACCTGG 1162
QY 1043 AGGAGCTGACAGAGGAACAGGCTCACCCACATCTTGAACATGAGCCCGGAGATTGACACT 1102
Db 1163 AGGAGCTGACAGAGGAACAGGCTCACCCACATCTTGAACATGAGCCCGGAGATTGACACT 1219
QY 1103 TCTACCTGAGCGCTTACCTTACCAACATGTGCGCTCTGGATGAGAGTCCGCCAGC 1162
Db 1220 TCTACCTGAGCGCTTACCTTACCAACATGTGCGCTCTGGATGAGAGTCCGCCAGC 1279
QY 1163 TGCTGCCGCACTGGAAGGAGCGCACCGCTTCAATTGAGGCTGCAAGAGCAGAGGCCACC 1222
Db 1280 TGCTGCCGCACTGGAAGGAGCGCACCGCTTCAATTGAGGCTGCAAGAGCAGAGGCCACC 1339
QY 1223 ACGTGTGCTCACTGCAAGATGGGCGTCAAGCGGCTCAGCGGCCACAGTGTGCTATG 1282
Db 1340 ACGTGTGCTCACTGCAAGATGGGCGTCAAGCGGCTCAGCGGCCACAGTGTGCTATG 1399
QY 1283 CCATGAGCAGTACGAATGAGCGCTGAGCAGGCGCTGCGCAAGTGAAGAGCTCCGGC 1342
Db 1400 CCATGAGCAGTACGAATGAGCGCTGAGCAGGCGCTGCGCAAGTGAAGAGCTCCGGC 1459
QY 1343 CCATGCGCGCGCCCAACCTGCTTCTGCGCAGCTGCAGATCTACAGGGCATCTGA 1402
Db 1460 CCATGCGCGCGCCCAACCTGCTTCTGCGCAGCTGCAGATCTACAGGGCATCTGA 1519
QY 1403 CGGCCAGCGCCAGAGCGCATGTCTGGAGCAGAAAGTGGTGGGTCTCCCAAGAGAGC 1462
Db 1520 CGGCCAGCGCCAGAGCGCATGTCTGGAGCAGAAAGTGGTGGGTCTCCCAAGAGAGC 1579
QY 1463 ACCAGCCCTGAAGTCTTACACCATCTTCCAGCTCTCCGCAAGAGAGGAGGTG 1522
Db 1580 ACCAGCCCTGAAGTCTTACACCATCTTCCAGCTCTCCGCAAGAGAGGAGGTG 1639
QY 1523 GGAAGGAGAGGTGTAGGATGGAAGAGAGGAGGAGCCCGGAGAGAGAGGCTGGGC 1582
Db 1640 GGAAGGAGAGGTGTAGGATGGAAGAGAGGAGGAGCCCGGAGAGAGGCTGGGC 1699

QY 1583 CACGGCCACGTATTAACCTCCGAGGGGTGATGAGTCCATCAGTCTTTGAGAGCCCTCCT 1642
DB 1700 CACGGCCACGTATTAACCTCCGAGGGGTGATGAGTCCATCAGTCTTTGAGAGCCCTCCT 1759
QY 1643 TGGAGCTGAGAGACACCTCCAGAGACCAGTACATGCGCAGAGGTCTTCTTCCACAGAGT 1702
DB 1760 TGGAGCTGAGAGACACCTCCAGAGACCAGTACATGCGCAGAGGTCTTCTTCCACAGAGT 1819
QY 1703 CTTACATGAGAGACCTCTGAGAGCCCTTCCCAAGAGTCCAGAGAGGAGGAGCCAGC 1762
DB 1820 CTTACATGAGAGACCTCTGAGAGCCCTTCCCAAGAGTCCAGAGAGGAGGAGCCAGC 1879
QY 1763 AGGTGAGAGAGGGGGCTCAGAGCTGAGAGTCCAGAGTCCAGAGTCCAGAGTCCAGAG 1822
DB 1880 AGGTGAGAGAGGGGGCTCAGAGCTGAGAGTCCAGAGTCCAGAGTCCAGAGTCCAGAG 1939
QY 1823 GCAGTGGCTGAGTGGCCAAACCCGAGCCCTTCCAGAGAGCAGAGAGAGGGGAGGGGC 1882
DB 1940 GCAGTGGCTGAGTGGCCAAACCCGAGCCCTTCCAGAGAGCAGAGAGAGGGGAGGGGC 1999
QY 1883 AGGGCAGAGAGAGCCCTGAGTCTTCTTACAGCCAGAGTCCAGAGAGTGGTGAAGACAG 1942
DB 2000 AGGGCAGAGAGAGCCCTGAGTCTTCTTACAGCCAGAGTCCAGAGAGTGGTGAAGACAG 2059
QY 1943 CCAGCGTGCATGACAGTGAAGAGAGGGCG 1972
DB 2060 CCAGCGTGCATGACAGTGAAGAGAGGGCG 2089

RESULT 7

US-09-955-732-20
; Sequence 20, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-955-732-20

Query Match 71.2%; Score 1409; DB 10; Length 1416;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCCCCTGTCACAGTGAAGCCGTTGCCCCCGGCGAGCGGCGCTCCACGCCCCGTGGGG 60
DB 1 ATGCCCCCTGTCACAGTGAAGCCGTTGCCCCCGGCGAGCGGCGCTCCACGCCCCGTGGGG 60
QY 61 CCCTGGAGCAGCGGGTCCAGCGAAGAGTGAAGTCCAGCGAAGGAGAGAGCTTTGGGGTG 120
DB 61 CCCTGGAGCAGCGGGTCCAGCGAAGAGTGAAGTCCAGCGAAGGAGAGAGCTTTGGGGTG 120
QY 121 CTCCGTGGGCTGTCTCTGGAGTGAAGTGAAGGGGCAATGATGATGACAGAGGGC 180
DB 121 CTCCGTGGGCTGTCTCTGGAGTGAAGTGAAGGGGCAATGATGATGACAGAGGGC 180
QY 181 AGTCTGAGCCACAAGAGAGGCGCCGAGTGAAGAGAGAGTCCAGCGGAGCAGACAGAC 240
DB 181 AGTCTGAGCCACAAGAGAGGCGCCGAGTGAAGAGAGAGTCCAGCGGAGCAGACAGAC 240
QY 241 TTGGGCAAGATCCAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 TTGGGCAAGATCCAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 ATGTACAGCTGCTGAGCGCGAGAGATGACATCCGCTGGCAGCCAGCTGAGAGAGAGAG 360

DB 301 ATGTACAGCTGCTGAGCGCGAGAGATGACATCCGCTGGCAGCCAGCTGAGAGAGAGAG 360
QY 361 CGGCTCCCGGGCTCCGCTACCTGCTGTAGTTCTTACACAGAGAGAGAGAGAGAGAGAG 420
DB 361 CGGCTCCCGGGCTCCGCTACCTGCTGTAGTTCTTACACAGAGAGAGAGAGAGAGAGAG 420
QY 421 CAGATGAGACCGGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 CAGATGAGACCGGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 CTGGGCTGGTCTTGGCCCTCTGAGAGTGAACCCAGAGTGTACTTAGATGAGAGAGGGGGC 540
DB 481 CTGGGCTGGTCTTGGCCCTCTGAGAGTGAACCCAGAGTGTACTTAGATGAGAGAGGGGGC 540
QY 541 TTCAGCGTGAAGTCTGAGTGGGCAAGCCGAGTCTTCAAGCCCATCTCCATCCAGACCATG 600
DB 541 TTCAGCGTGAAGTCTGAGTGGGCAAGCCGAGTCTTCAAGCCCATCTCCATCCAGACCATG 600
QY 601 TGGGCACTCCAGGTATGACCAAGCATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 TGGGCACTCCAGGTATGACCAAGCATGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CCGGCTGAGAGTCCCTCAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 CCGGCTGAGAGTCCCTCAGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 AGCTGCTCAATGAGTGAAGCGGTATGAGCGAGTGTCTGCGGCTCCAGAGCGCC 780
DB 721 AGCTGCTCAATGAGTGAAGCGGTATGAGCGAGTGTCTGCGGCTCCAGAGCGCC 780
QY 781 GAGCTGGCGGGTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 GAGCTGGCGGGTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 AAAGTGTGATGTCAAGTGAAGAGTGTCACTTCCAAAGAGATCCGCGAGCTCTG 900
DB 841 AAAGTGTGATGTCAAGTGAAGAGTGTCACTTCCAAAGAGATCCGCGAGCTCTG 900
QY 901 GAGCTGGCGGGTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB 901 GAGCTGGCGGGTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CTGCTGTGAGCAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 CTGCTGTGAGCAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 GAGTGAAGCAGCAAACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 GAGTGAAGCAGCAAACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 ATGCCCCGGGAGATTGACAACTTCACTGAGCGCTTCACTACCAATGTGCGCTC 1140
DB 1081 ATGCCCCGGGAGATTGACAACTTCACTGAGCGCTTCACTACCAATGTGCGCTC 1140
QY 1141 TGGATGAGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 TGGATGAGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GCTGCAAGAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 GCTGCAAGAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 GCGGCAAGAGTGGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 GCGGCAAGAGTGGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 CGCCAGTGAAGAGTCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1321 CGCCAGTGAAGAGTCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 CAGATCTACAGGAGATCTGAGCGGAG 1409

Db 1381 CAGATCTACCAGGCGCATCCTGACGGCCAG 1409

RESULT 8

US-09-761-640-3
; Sequence 3, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C1000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2540
; TYPE: DNA
; ORGANISM: Human
US-09-761-640-3

Query Match 68.0%; Score 1345.6; DB 10; Length 2540;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 4; Indels 291; Gaps 5;

QY 1 ATGGCCCTGTACAGTAGCCGTTCCGCCCGGCGAGCGCGCTCCACGCCCGTGGG 60
Db 88 ATGGCCCTGTACAGTAGCCGTTCCGCCCGGCGAGCGCGCTCCACGCCCGTGGG 147
QY 61 CCCTGGGACCGGCGGTCCAGCGAAGAGTCACTCCAGCGAAGCAGAGCTTTGCGGTG 120
Db 148 CCCTGGGACCGGCGGTCCAGCGAAGAGTCACTCCAGCGAAGCAGAGCTTTGCGGTG 207
QY 121 CTCCGTGGGGCTGTCTTGGGACTGCAGGATGAGGGGCAATGATGTCAGCAGAGGCC 180
Db 208 CTCCGTGGGGCTGTCTTGGGACTGCAGGATGAGGGGCAATGATGTCAGCAGAGGCC 267
QY 181 AGTTCTGAGCCAAACAGAGAGGCCCGGAGTGAAGAGAGCTCCAGGGGACCAAGACAG 240
Db 268 AGTTCTGAGCCAAACAGAGAGGCCCGGAGTGAAGAGAGCTCCAGGGGACCAAGACAG 327
QY 241 TTGGGGCAAGGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 328 TTGGGGCAAGGATCCAGAGTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 387
QY 301 ATGTACAGCTGTGAGGCGGAGATGATCCGCTGCGAGCCAGCTGGAGGCAACC 360
Db 388 ATGTACAGCTGTGAGGCGGAGATGATCCGCTGCGAGCCAGCTGGAGGCAACC 447
QY 361 CGGCTCCCGGCTCCGCTACCTGTGTGATTCTTACACGAGAGAGAGAGAGTCTGAGC 420
Db 448 CGGCTCCCGGCTCCGCTACCTGTGTGATTCTTACACGAGAGAGAGAGTCTGAGC 507
QY 421 CAGGATGAGACGCTCTCTGGGCGTGGATTTCCTGACAGCAGCTCCCCAGCTGACCC 480
Db 508 CAGGATGAGACGCTCTCTGGGCGTGGATTTCCTGACAGCAGCTCCCCAGCTGACCC 567
QY 481 CTGGGCTGTGTTGCCCTCTGAGTGAACCCAGGTGTACTAGATGAGACGGGGGC 540
Db 568 CTGGGCTGTGTTGCCCTCTGAGTGAACCCAGGTGTACTAGATGAGACGGGGGC 627
QY 541 TTCAAGCTGACGTCTGTGGGCAAGCGCGGATCTTCAAGCCCATCTCCATCCAGACCA 600
Db 628 TTCAAGCTGACGTCTGTGGGCAAGCGCGGATCTTCAAGCCCATCTCCATCCAGACCA 687
QY 601 TGGGCCACACTCCAGGTAATTGCAACCAAGCATGTAGGCGAGCTTAAGGCAAGCGGCTT 660
Db 688 T----- 688
QY 661 CCGGGTGGCAGTGCCTCACTGGGCGAGCCACTACAGAGAGAGACTGAACCTCCGAACAG 720

Db 689 ----- 688
QY 721 AGCTGCTCAATGAGTGAAGGCTATGGCCGACTGGAGTCTCTGGGCTCCAGCGCC 780
Db 689 ----- 688
QY 781 GAGCCTGGCGGCTCTCAGAACAGAGACAGATGAGACAGCGGATCCGTGTAGCTGTG 840
Db 689 -----GCTCCTCAGAACAGAGACAGATGAGACAGCGGATCCGTGTAGCTGTG 738
QY 841 AAGTGTTCATGTCAGTGAACCTGAGAGTGTCTACTTCCAAAGAGATCCGCCAGGCTCTG 900
Db 739 AAGTGTTCATGTCAGTGAACCTGAGAGTGTCTACTTCCAAAGAGATCCGCCAGGCTCTG 798
QY 901 GAGCTGGGCTGGGGCTCCCTCCAGCAGTACCGTCACTTCATGACAAACAGATGCTG 960
Db 799 GAGCTGGGCTGGGGCTCCCTCCAGCAGTACCGTCACTTCATGACAAACAGATGCTG 858
QY 961 CTGCTGTGGCACAGCGGACCGAGCCTCCCGCATCTTCCCGACCTTCACTGGGCTCA 1020
Db 859 CTGCTGTGGCACAGCGGACCGAGCCTCCCGCATCTTCCCGACCTTCACTGGGCTCA 918
QY 1021 GAGTGAACGACGAAACCTGAGAGAGCTGCAGAGAAACAGGGTCAACCCACATCTTGAAC 1080
Db 919 GAGTGAACGACGAAACCTGAGAGAGCTGCAGAGAAACAGGGTCAACCCACATCTTGAAC 978
QY 1081 ATGGCCGGGAGATTGACAACTTCTACCTGAGGCTTCACTACCAATGTGCGCCTC 1140
Db 979 ATGGCCGGGAGATTGACAACTTCTACCTGAGGCTTCACTACCAATGTGCGCCTC 1038
QY 1141 TGGATGAGAGATCGGCCAGCTGCTGCGCACTGGAAGAGACGACCGCTTATTGAG 1200
Db 1039 TGGATGAGAGATCGGCCAGCTGCTGCGCACTGGAAGAGACGACCGCTTATTGAG 1098
QY 1201 GCTGCAAGACACAGAGGCAACCAAGTGTGCTGCTCACTGCAAGATGGGCGTCAAGCCGCTCA 1260
Db 1099 GCTGCAAGACACAGAGGCAACCAAGTGTGCTGCTCACTGCAAGATGGGCGTCAAGCCGCTCA 1158
QY 1261 GCGGCACAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACAGCTTGAGAGCGCCCTG 1320
Db 1159 GCGGCACAGTGTGCTGCTATGCCATGAAGCAGTACGAATGACAGCTTGAGAGCGCCCTG 1218
QY 1321 CGCCACGTGACGAGGCTCCGCGCCCATGCGCCCGCCCAACCTGGCTTCTGCGCCAGCTG 1380
Db 1219 CGCCACGTGACGAGGCTCCGCGCCCATGCGCCCGCCCAACCTGGCTTCTGCGCCAGCTG 1278
QY 1381 CAGATTAACAGGCGCATCTGACGCGCCAGCGCCGAGAGCCATGTCTGGAGACAGAAAGTG 1440
Db 1279 CAGATTAACAGGCGCATCTGACG----- 1302
QY 1441 GGTGGGCTTCCCGAGAGAGACCCAGCCCTGAAGTCTTACACCAATTCCCACTCTT 1500
Db 1303 ----- 1302
QY 1501 CCGCAGAACTGAGGGTGTGGGAGAGAGAGTGTAGGCATGAAAGAGAGCCAGGCA 1560
Db 1303 --GCCAGAACTGAGGGTGTGGGAGAGAGAGTGTAGGCATGAAAGAGAGCCAGGCA 1360
QY 1561 GCGCCGAAAAGAGAGCTGGG--CCACGGCCAGTATAAACCTCCGAGGGGTATGAGGT 1618
Db 1361 GCGCCGAAAAGAGAGCTGGGGCCACGGGGCAGGTATAAACCTCCGAGGGGTATGAGGT 1420
QY 1619 CCATCAGTCTTCTGAGGCTCTCT--GAGCTGAGAGACACCTCAG--AGACAGTGACAT 1676
Db 1421 CCATCAGTCTTCTGAGGCTCTCTTGGAGCTGAGAGACACCTCAGTAGACATGACAT 1480
QY 1677 GCCAGAGTCTTCTCTTCCACAGAGTCTTCAATGAAGAGCTCTGAGCCCTTCCACA 1736
Db 1481 GCCAGAGTCTTCTCTTCCACAGAGTCTTCAATGAAGAGCTCTGAGCCCTTCCACA 1540
QY 1737 GCTTGAAGAGACCAAGGAGGCGACAGAGTGAAGAGGGGCTGAGCTGCGCTGAAGTC 1796
Db 1541 GCTTGAAGAGACCAAGGAGGCGACAGAGTGAAGAGGGGCTGAGCTGCGCTGAAGTC 1600


```
RESULT 10
US-10-363-676-4
; Sequence 4, Application US/10363676
; Publication No. US20030170856A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN MAP KINASE PHOSPHATASE-LIKE ENZYME
; FILE REFERENCE: L10122 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/363,676
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US 60/230,709
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-363-676-4

Query Match      35.1%; Score 694; DB 13; Length 1755;
Best Local Similarity 89.1%; Pred. No. 1.9e-171;
Matches 802; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

QY      1081 ATGCCCCGGAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCAATGTGCGCTC 1140
Db      322 ATGGCCCCGGAGATTGACAACTTCTACCTTGAGCGCTTCACTTACCAATGTGCGCTC 381

QY      1141 TGGGATGAGAGTGGGCCCACTGCTGCTGCCGCACTGGAAGAGAGCAACCGCTTCATTGAG 1200
Db      382 TGGGATGAGAGTGGGCCCACTGCTGCTGCCGCACTGGAAGAGAGCAACCGCTTCATTGAG 441

QY      1201 GCTGCAAGAGCACAGGGGCAACCCACGTGCTGCTGCTCACTGCAAGATGGGCGTCAGCCGCTCA 1260
Db      442 GCTGCAAGAGCACAGGGGCAACCCACGTGCTGCTGCTCACTGCAAGATGGGCGTCAGCCGCTCA 501

QY      1261 GCGGCCACAGTGTGCTGCTATGCCATGAAGAGATGCAATGCAAGCTTGAGCAGGCGCTG 1320
Db      502 GCGGCCACAGTGTGCTGCTATGCCATGAAGAGATGCAATGCAAGCTTGAGCAGGCGCTG 561

QY      1321 GCGCCAGTGCAGAGCTCCGGCCCATGCGCCGCCCAACCCCTGCTTCTCGGCCAGCTG 1380
Db      562 GCGCCAGTGCAGAGCTCCGGCCCATGCGCCGCCCAACCCCTGCTTCTCGGCCAGCTG 621

QY      1381 CAGATCTACCAAGGCACTCCAGCGGCCAGCGCCAGAGCCATGCTGGAGCAGAAAGTG 1440
Db      622 CAGATCTACCAAGGCACTCCAGCGGCCAGCGCCAGAGCCATGCTGGAGCAGAAAGTG 645

QY      1441 GGTGGGGTCTCCCAAGAGAGCAACCCAGCCCTGAAGTCTCTACACCATTCCCACTCTT 1500
Db      646 ----- 645

QY      1501 CCGCCAGAACCTGAGGGTGTGGGGAGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCA 1560
Db      646 --GCCAGAACCTGAGGGTGTGGGGAGAGAAAGTTGTAGGCATGGAAGAGAGCCAGGCA 703

QY      1561 GCCCCGAAAGAGAGCTGGGCCACGGGCCAGTATAACCTCCGAGGGGTCTAGAGGTCC 1620
Db      704 GCCCCGAAAGAGAGCTGGGCCACGGGCCAGTATAACCTCCGAGGGGTCTAGAGGTCC 763

QY      1621 ATCAGTCTTGTGAGCCCTCTTGAGCTGGAGAGCACCCTCAGAGACCAGTGACATGCCA 1680
Db      764 ATCAGTCTTGTGAGCCCTCTTGAGCTGGAGAGCACCCTCAGAGACCAGTGACATGCCA 823

QY      1681 GAGGTCTTCTTCCACGAGTCTTACATGAAGAGCCTCTGAGCCCTTCCACAGCTT 1740
Db      824 GAGGTCTTCTTCCACGAGTCTTACATGAAGAGCCTCTGAGCCCTTCCACAGCTT 883

QY      1741 GCAAGGACCAAGGGAGGCGCAGAGTGGAGAGGGGGCTCAGCCTGCGCTGAAGTCCGC 1800
Db      884 GCAAGGACCAAGGGAGGCGCAGAGTGGAGAGGGGGCTCAGCCTGCGCTGAAGTCCGC 943
```

```
QY      1801 CAGTCAGTGGTTACCTCCAGGGCAGTGCCTGTGTGCCAACCCGAGCCAGGCTTCCAG 1860
Db      944 CAGTCAGTGGTTACCTCCAGGGCAGTGCCTGTGTGCCAACCCGAGCCAGGCTTCCAG 1003

QY      1861 GAGCAGAGCAGGGGCGAGGGGCGAGGGAGAGAGCCCTGCATTCTCTACGCCAGG 1920
Db      1004 GAGCAGAGCAGGGGCGAGGGGCGAGGGAGAGAGCCCTGCATTCTCTACGCCAGG 1063

QY      1921 TTCCGAAGGTGTGAGACAGGCCACGCTGCATGACAGTGAGAGAGAGGGCGAGCCTGA 1980
Db      1064 TTCCGAAGGTGTGAGACAGGCCACGCTGCATGACAGTGAGAGAGAGGGCGAGCCTGA 1123

RESULT 11
US-10-363-676-8
; Sequence 8, Application US/10363676
; Publication No. US20030170856A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN MAP KINASE PHOSPHATASE-LIKE ENZYME
; FILE REFERENCE: L10122 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/363,676
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US 60/230,709
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 599
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-363-676-8

Query Match      24.8%; Score 491; DB 13; Length 599;
Best Local Similarity 85.9%; Pred. No. 1.9e-118;
Matches 599; Conservative 0; Mismatches 0; Indels 98; Gaps 1;

QY      1024 TGAACGACGAAACCTGAGGAGCTGCAGAGAAAGGGTCAACCCATCTGAACATG 1083
Db      1 TGAACGACGAAACCTGAGGAGCTGCAGAGAAAGGGTCAACCCATCTGAACATG 60

QY      1084 GCGCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATGTGCGCTCTGG 1143
Db      61 GCGCGGAGATTGACAACTTCTACCTGAGCGCTTCACTACCAATGTGCGCTCTGG 120

QY      1144 GATGAGAGTGGGCCAGCTGCTGCCGCACTGGAAGAGAGCAGCGCTTCATTGAGGCT 1203
Db      121 GATGAGAGTGGGCCAGCTGCTGCCGCACTGGAAGAGAGCAGCGCTTCATTGAGGCT 180

QY      1204 GCAAGAGCACAGGGCACCCACGTGCTGTGTCATGCAAGATGGGGCTCAGCCGCTCAGCG 1263
Db      181 GCAAGAGCACAGGGCACCCACGTGCTGTGTCATGCAAGATGGGGCTCAGCG 240

QY      1264 GCCACAGTGTGGCTATGCCATGAAGCAGTACGAATGCAGCCCTGAGCAGGCCCTGCGC 1323
Db      241 GCCACAGTGTGGCTATGCCATGAAGCAGTACGAATGCAGCCCTGAGCAGGCCCTGCGC 300

QY      1324 CACGTGCAGGAGCTCCGGCCCATGCGCCGCCCAACCCCTGGCTTCTGCGCCAGCTGCAG 1383
Db      301 CACGTGCAGGAGCTCCGGCCCATGCGCCGCCCAACCCCTGGCTTCTGCGCCAGCTGCAG 360

QY      1384 ATCTACAGGGCATCTTAACGGCCAGCGCCAGAGCCATGTCTGGAGCAGAAAGTGGGT 1443
Db      361 ATCTACAGGGCATCTTAACG----- 381

QY      1444 GGGGTCTCCCAAGAGAGCAACCCCTGAAGTCTCTACACCAATTCCCACTCTTCCG 1503
Db      382 -----G 382

QY      1504 CCAGAACCTGAGGGTGTGGGGAGGAGAAAGTTGTAGGCATGGAAGAGCCAGGAGCC 1563
Db      383 CCAGAACCTGAGGGTGTGGGGAGGAGAAAGTTGTAGGCATGGAAGAGCCAGGAGCC 442
```


Db 1156 AATCTGAGGAAGTCTGAGGAGCTCAGGGGTGATTTACATTTTAATGTATACCAGAGAAATC 1215
QY 1096 GACAACTTCTACCTGAGCGCTTACCTTACCAATGTGCGCTCTGGGATGAGGATCG 1155
Db 1216 GATTAATTTTTCCTGGCTTATTTGATATATACATCCGAGTCTACGATGAAGAGACC 1275
QY 1156 GCGCAGCTGCTGCGCACTGGAAGAGAGACGACCGCTTCAATTGAGGCTGCAAGAGACAG 1215
Db 1276 ACAGACCTCTCGCCCACTGGAATGAAGCGTATCATTTTATAACAAGCGAAGAGAGAAC 1335
QY 1216 GGCACCCACGCTGCTGTCCACTGCAAGATGGCGCTCAGCCGCTCAGCGGCCACAGTGTG 1275
Db 1336 CATTCAGAGTGCCTGTGTCATTGCAAAATGGCGCTGAGTGCCTGCGCTCCACAGTCATA 1395
QY 1276 GCGTATGCCATGAAGCAGTACGATGACGCTGAGCAGAGCCCTGCGCCACGTCAGAGAG 1335
Db 1396 GCGTATGCAATGAAGGAAATTCGGCTGCGCTGTGAAAAAAGCATATACTATGTAAGCAG 1455
QY 1336 CTCGGGCGCATGCGCGCGCCCAACCCCTGCTTCTGCGCCAGCTGCAATCTACCAAGGC 1395
Db 1456 AAGCGCAGCATCAGCGCGCCCAACGCGGCTTATGAGGACAGCTGTCTGAGTATGAAGGC 1515
QY 1396 ATCTTGACGCGCCAGCGCCAGAGCCA 1421
Db 1516 ATCTTGATGCAAGCAAAACAGCGGCA 1541

RESULT 14

US-10-363-676-9
; Sequence 9, Application US/10363676
; Publication No. US20030170856A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN MAP KINASE PHOSPHATASE-LIKE ENZYME
; FILE REFERENCE: L10122 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/363,676
; PRIOR FILING DATE: 2003-03-06
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-363-676-9

Query Match 18.0%; Score 356.4; DB 13; Length 494;
Best Local Similarity 98.4%; Pred. No. 2.7e-83;
Matches 360; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1169 CGCACTGGAAGAGAGACGCAACCGCTTCAATTGAGGCTGCAAGAGACAGGGCACCCACGCTGC 1228
Db 1 CGCACTGGAAGAGAGACGCAACCGCTTCAATTGAGGCTGCAAGAGACAGGGCACCCACGCTGC 60
QY 1229 TGGTCCACTGCAAGATGGGCGTACGCGCTCAGCGGCCACAGTGTGCTGCTATGTCATGA 1288
Db 61 TGGTCCACTGCAAGATGGGCGTACGCGCTCAGCGGCCACAGTGTGCTGCTATGTCATGA 120
QY 1289 AGCAGTACGAATGAGCCTGAGCAGAGGCGCTGCGCCACGTCGAGAGAGCTCCGGCCATCG 1348
Db 121 AGCAGTACGAATGAGCCTGAGCAGAGGCGCTGCGCCACGTCGAGAGAGCTCCGGCCATCG 180
QY 1349 CCGCGCCCAACCTGGCTTCTGCGCCAGCTGCAAGTCTACAGGGCATCTGAAGGCCA 1408
Db 181 CCGCGCCCAACCTGGCTTCTGCGCCAGCTGCAAGTCTACAGGGCATCTGAAGGCCA 240
QY 1409 GCGCGCAGAGCCATGTCTGGAGCAGAAAGTGGGTGGGTCTCCAGAGAGACCCAG 1468
Db 241 GCGCGCAGAGCCATGTCTGGAGCAGAAAGTGGGTGGGTCTCCAGAGAGACCCAG 300
QY 1469 CCCCTGAAGTCTCTACACCATTTCCCACTCTTCCGCCAAGACCTGAGGCTGTTGGGAGG 1528

Db 301 CCCCTGAAGTCTCTACACCATTTCCCACTCTTCCGCCAAGACCTGAGGCTGGTGGCTCG 360
QY 1529 AGAAG 1534
Db 361 GGGGG 366

RESULT 15

US-09-775-925-1
; Sequence 1, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; FILE REFERENCE: PHOSPHATASES
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-775-925-1

Query Match 17.2%; Score 340.4; DB 9; Length 1949;
Best Local Similarity 59.5%; Pred. No. 5.1e-79;
Matches 597; Conservative 0; Mismatches 401; Indels 6; Gaps 1;

QY 424 GATGAGACGCTCTCTCTGGCGGTGATTTCCCTGACAGCAGCTCCCCAGCTGCACCCTG 483
Db 32 GAGAGATATATCTTGTCTGGAGTGAAGTCTTCCAGTAAGAAAGTAAAGCTGCACCAT 91
QY 484 GGCTGTCTTGCCTCTGAGTGAACCCAGGTGTACTTAGATGAGACGGGGCTTC 543
Db 92 GGGATGTTCTCCGACTGTGAGCGACGCAAAATCCACCTTGATGAGATGTGGTTC 151
QY 544 AGCGTACGCTGTGTGGCAAGCCGGATCTTCAAGCCCATCTCCATCCAGACCATGTG 603
Db 152 AGCGTACGACAGCAGGAAGGATGACATATTTAAGCCTGTGTCTCCAGGCCATGTG 211
QY 604 GCCACACTCCAGGTATTCACCAAGCATGTAGGACGCTTAGGCGCGCTGTACCG 663
Db 212 TCTGCCCTGCAAGTGTCTTCAAGGCTGCGAAGTGGCCCGGAGGCAACAATCTTCCC 271
QY 664 GGTGGCAGTGCCTCACTGCGCCAGCCACTACAGAGAGACTGAACCTCGAAGAGAG 723
Db 272 GGGGTGTAGCTCTCACTGTGGCTACCTATAGAGCTGCATCACTCCGAGCAGAGC 331
QY 724 TGCTCAATGATGAGCGGCTATAGCCGACTGAGTCTTGGCGCTCCAGCGCGAG 783
Db 332 TGCAATCAAGAGTGAAGCCCATGACAGACTGAGTCTACGCGCGCCGACTCCCGCG 391
QY 784 C-----CTGGCGGTCTCAGAAACAGAGCAGATGAGCAGGCGATCCGTGAGCTG 837
Db 392 CTATTTGTGACAAAGCCCACTGAAGGGGAAAGAACCGAGCGCTCATCAAGCCAGCTC 451
QY 838 TGAAGAAGTGTGATGTCACTGAGAGTGTCACTTCCAAAGATCCGCGAGCT 897
Db 452 CGAAGCATCATGATGAGCCAGATCTGAAGAAATGTGACTTCCAAAGATTCGTAATGA 511
QY 898 CTGAGCTGCGCTGCGGCTCCCTCCAGCAGTACCGTGAATCATGCAACAACAGATG 957
Db 512 TTAGAGAAACAGATGAATTTGTAATTTGAAGAACTCAAGGAATTTATAGCAATGAGATG 571
QY 958 CTGCTGTGTGAGCAGCGGAGCGAGCTCCCGCATCTTCCCACTTACCTGGC 1017
Db 572 CTACTTATCTTGGACAGATGAGCAAGCCCTCCCTATCTTGAATCATCTTATCTGGC 631
QY 1018 TCAGAGTGAACGCAACCACTGAGAGCTGACAGAGAACGGTCAACCCACATCTTG 1077

Db 632 TCTGAATGGAATGCATCAATCTGAGGAAGCTGACAGGGCTCAGGGGTGATTACATTTTA 691
QY 1078 AACATGGCCCGGAGATTGACAACTTCTTACCTGAGCGCTTACCTACCAATGTGGC 1137
Db 692 AATGTTACAGAGAAATGATAATTTTCTTGGCTTATTGATATCATPAACATCCGA 751
QY 1138 CTCTGGATGAGAGTGGCCAGCTGTGCTGCGCACTGGAAGGAGACGCAACGCTTCATT 1197
Db 752 GTCTACGATGAAGAGACCAAGACCTCCTGCGCACTGGAATGAAGCTATCATTTTATA 811
QY 1198 GAGCTGCAAGAGACACAGGCAACCAAGTGTGCTGCTCAAGATGGCGTCAGCGGC 1257
Db 812 AACAAAGCGAAGAGAACCAATTCAGAGTGGCTGTGCAATTGCAAAATGGCGTGAGTCGC 871
QY 1258 TCAGCGGCCACAGTGTGGCTTATGCCATGAAGCAGTACGAATGCAAGCTTGAGCAGGCC 1317
Db 872 TCGGCTCCACAGTCATAGCCTATGCAATGAAGAAATTCGCTGGCTCTGAAAAAGCA 931
QY 1318 CTGCGCCACGTGCAGAGGCTCCGGCCCATCGCCGCCCAACCTGGCTTCTGCGCCAG 1377
Db 932 TATACTATGTAAGCAGAGCGCAGCATCACGGGCCCAACGGGCTTTATGAGGCAG 991
QY 1378 CTGCAATCTACCAAGGCGCATCTGACGCGCCAGCGCCAGAGCCA 1421
Db 992 CTGTCTAGTATGAAGGCACTTGGATGCAAGCAACAGCGGCA 1035

Search completed: January 15, 2004, 20:27:08
Job time : 671 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2004, 15:07:05 ; Search time 143 Seconds
(without alignments)
611.459 Million cell updates/sec

Title: US-09-955-732-1
Perfect score: 1980
Sequence: 1 atggccctgtcacagtgag.....gagaggagggcgagcctga 1980

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101.8	5.1	531	US-09-704-139-3	Sequence 3, Appli
2	101.8	5.1	1390	US-09-704-139-1	Sequence 1, Appli
3	85.4	4.3	1238	US-08-530-290-11	Sequence 11, Appli
4	85.4	4.3	1238	US-09-702-705-803	Sequence 803, App
5	85.4	4.3	1238	US-09-736-457-803	Sequence 803, App
6	85.4	4.3	2064	US-09-702-705-825	Sequence 825, App
7	85.4	4.3	2064	US-09-736-457-825	Sequence 825, App
8	85.4	4.3	2109	US-09-702-705-826	Sequence 826, App
9	85.4	4.3	2109	US-09-736-457-826	Sequence 826, App
10	85.4	4.3	2240	US-09-016-434-1100	Sequence 1100, Ap
11	84	4.2	2303	US-09-922-146-3	Sequence 3, Appli
12	83.8	4.2	1619	US-09-702-705-801	Sequence 801, App
13	83.8	4.2	1619	US-09-736-457-801	Sequence 801, App
14	83.8	4.2	4637	US-09-702-705-804	Sequence 804, App
15	83.8	4.2	4637	US-09-736-457-804	Sequence 804, App
16	83	4.2	247	US-09-016-434-864	Sequence 864, App
17	81.4	4.1	2377	US-09-920-668-3	Sequence 3, Appli
18	77.4	3.9	1987	US-08-990-379-1	Sequence 1, Appli
19	77.4	3.9	1993	US-08-990-379-2	Sequence 2, Appli
20	73.6	3.7	1830	US-09-557-921-1	Sequence 1, Appli
21	71.2	3.6	2109	US-09-016-434-1135	Sequence 1135, Ap
22	70.6	3.6	2000	US-09-016-434-1291	Sequence 1291, Ap
23	68.4	3.5	944	US-09-371-671B-10	Sequence 10, Appli
24	65.6	3.3	539	US-09-389-681-311	Sequence 311, App
25	65.6	3.3	539	US-09-620-405B-311	Sequence 311, App
26	65.6	3.3	539	US-09-339-338-311	Sequence 311, App
27	65.6	3.3	539	US-09-433-826B-311	Sequence 311, App

C	28	65.6	3.3	539	4	US-09-604-287A-311	Sequence 311, App
	29	65	3.3	2450	4	US-09-770-595A-21	Sequence 21, Appl
	30	61.6	3.1	1729	3	US-09-045-973-6	Sequence 6, Appli
	31	56.2	2.8	594	3	US-09-163-833-3	Sequence 3, Appli
	32	56.2	2.8	912	3	US-09-163-833-1	Sequence 1, Appli
	33	55.8	2.8	1227	4	US-09-620-312D-1027	Sequence 1027, Ap
	34	54.6	2.8	1380	4	US-09-620-312D-1026	Sequence 1026, Ap
	35	53.2	2.7	1691	3	US-09-013-881-12	Sequence 12, Appl
	36	53.2	2.7	1691	4	US-09-612-473-12	Sequence 12, Appl
	37	53.2	2.7	1878	4	US-09-620-312D-398	Sequence 398, App
	38	52.6	2.7	216	2	US-08-530-290-9	Sequence 9, Appli
	39	51.6	2.6	1491	4	US-09-544-716-1	Sequence 1, Appli
	40	50.2	2.5	861	1	US-07-988-273-1	Sequence 1, Appli
	41	50.2	2.5	861	5	PCT-US93-12019-1	Sequence 1, Appli
	42	48.6	2.5	477	3	US-09-135-994-1	Sequence 1, Appli
	43	48.6	2.5	477	4	US-09-684-843A-1	Sequence 1, Appli
	44	48.2	2.4	237	2	US-08-530-290-8	Sequence 8, Appli
C	45	48	2.4	3889	4	US-09-484-970B-39	Sequence 39, Appli

ALIGNMENTS

RESULT 1									
US-09-704-139-3									
; Sequence 3, Application US/09704139									
; Patent No. 6420153									
; GENERAL INFORMATION:									
; APPLICANT: Kapeller, Rosanna									
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR									
; FILE REFERENCE: 10448-018001									
; CURRENT APPLICATION NUMBER: US/09/704, 139									
; CURRENT FILING DATE: 2000-11-01									
; PRIOR APPLICATION NUMBER: US 60/185, 772									
; PRIOR FILING DATE: 2000-02-29									
; NUMBER OF SEQ ID NOS: 5									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 3									
; LENGTH: 531									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; US-09-704-139-3									
Query Match									
Best Local Similarity 5.1%; Score 101.8; DB 4; Length 531;									
Matches 220; Conservative 0; Mismatches 197; Indels 0; Gaps 0;									
QY	991	CGCATCTTCCCCACCTTCACTGCTGAGTGGAACGAGCAACTGAGAGCTG	1050						
DB	61	CGCGTCGCGCCCTCACTCTTCTCGGAGCGCGGAGCGCGGAGAGAGCTG	120						
QY	1051	CAGAGAACAGGTCACCAATCTTGAACATGCGCGGAGATTGACACTTACCT	1110						
DB	121	GCGCGCGGAGTACAGCTGTGCTCAACGTCTCCGACAGACCGCGCGCGCG	180						
QY	1111	GAGCGTTCACTTACCAATGTGCGCTCTGGATGAGAGTGGCCCACTGCTGCG	1170						
DB	181	CCCGGCTGGCAGAGCTGCGCGTCCGCTTTCAGACGCCGCTGAGAGACTGCTG	240						
QY	1171	CACTGAAGAGAGACGACCGCTTCACTTGAAGGTCAGAGACAGAGGACCCACG	1230						
DB	241	CACCTGAGAGCCACGTCGCGCGCATGAGGCGCGGCGCGCGCGCGCTGCTA	300						
QY	1231	GTCACTGCAAGATGGGCTTCAAGCGCTCAAGCGGACAGAGTGTGCTATGCTATG	1290						
DB	301	GTCTACTGCAAGACGCGCGGACCGCTCGCGCGCTGCAAGGCTATGCTATGCT	360						
QY	1291	CAGTACGAATGACGCTGAGAGAGCGCTGCGGACAGTGCAGAGAGCTTCGGCC	1350						
DB	361	CACCGCGGCTTCAAGCTTGGCGGAGGCTTCCAGATGTGAAGAGCGCTCGCC	420						
QY	1351	CGCCCAACCTTGCTTCTGCGGAGCTGCAATCTACAGGAGGCTCTGACGCGC	1407						

DB 421 GAACCGAACCCGGGCTTCTGCTCTCAGCTCCAGAGATAGAGAGGCCCTCCAGGCC 477

RESULT 2

US-09-704-139-1

Sequence 1, Application US/09704139

Patent No. 6420153

GENERAL INFORMATION:

APPLICANT: Kapeller, Rosanna

TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR

FILE REFERENCE: 10448-018001

CURRENT APPLICATION NUMBER: US/09/704,139

PRIOR FILING DATE: 2000-11-01

PRIOR APPLICATION NUMBER: US 60/185,772

PRIOR FILING DATE: 2000-02-29

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 1390

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (329)..(859)

NAME/KEY: misc feature

LOCATION: (1)..(1390)

OTHER INFORMATION: n = a, t, g, or c

US-09-704-139-1

Query Match

Best Local Similarity 5.1%; Score 101.8; DB 4; Length 1390;

Matches 220; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 991 CGCATCTTCCCGCACCTCTACCTGGGCTCAGAGTGAACGCAAACTGGAGGAGCTG 1050

DB 389 CGCGTGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 448

QY 1051 CAGAGGAACAGAGGTCACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCTAC 1110

DB 449 GCGCGCGGAGATCAGCTGTGCTGCTCAACGCTCTCCCGCAGCAGCCCGCGCGCG 508

QY 1111 GAGCGCTTCACTACCAATGTGGGCTCTGGATGAGAGTGGCCCACTGCTGCGG 1170

DB 509 CCGCGCTGGCAGAGCTGCGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 568

QY 1171 CACTGGAAGAGACGACCGCTTATGAGGCTGCAAGAGACACAGGACCACTGCTG 1230

DB 569 CACTGGAAGACGACCGCTTATGAGGCTGCAAGAGACACAGGACCACTGCTGCTG 628

QY 1231 GTCCACTGCAAGATGGGCTCAGCGGCTCAGCGGCAAGTGTGCTTATGCTATGAG 1290

DB 629 GTCTACTGCAAGAGAGCGGCGCAGCGGCTGCGCGGCTGCAACCGCTATCTATGCG 688

QY 1291 CAGTAGAATGACGCTTGAGCAGAGCGGCTGCGGCAAGTGTGCTTATGCTATGAG 1350

DB 689 CACCGCGGCTGAGCTTGCGGAGGCTTCCAGATGTGAAGAGCGCTGCGCGGCTAGCA 748

QY 1351 CGCCCCAACCTGGCTTCTGCGGCACTGCAAGTGTGCTTATGCTATGAGGCTGAG 1407

DB 749 GAACCGAACCCGGGCTTCTGCTCTCAGCTCCAGAGATAGAGAGGCCCTCCAGGCC 805

RESULT 3

US-08-530-290-11

Sequence 11, Application US/08530290

Patent No. 5958721

GENERAL INFORMATION:

APPLICANT: Marshall, Christopher John

APPLICANT: Ashworth, Alan

APPLICANT: Hughes, David Anthony

TITLE OF INVENTION: Methods for Screening of Substances for

TITLE OF INVENTION: Therapeutic Activity and Yeast for Use Therein

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/530,290

FILING DATE: 14-DEC-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/GB94/00694

FILING DATE: 31-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9402573.1

FILING DATE: 10-FEB-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9307250.2

FILING DATE: 07-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 084611-000000US

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 1238 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-530-290-11

Query Match

Best Local Similarity 4.3%; Score 85.4; DB 2; Length 1238;

Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCGCACCTCTACCTGGGCTCAGAGTGAACGCAAACTGGAGAGCTGAC 1053

DB 648 ATCTTCCCGCACCTCTACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 707

QY 1054 AGGAACAGGTCACCCACATCTTGAACATGGCCCGGAGATTGACAACTTCTAC 1113

DB 708 GCGCTGGGATCAGCGGCTCTGTTGATGTCTCTCTGAGTCCCAACCACTTTGAAGA 767

QY 1114 CGCTTCACTACCAATGTGGGCTCTGGATGAGAGTGGCCCACTGCTGCTGCGCAC 1173

DB 768 CACTATCAGTCAAGTGCATCCAGTGAAGATTAACCAAGGCGGACATCATCTCTGG 827

QY 1174 TGAAGAGAGACGACCGCTTATGAGGCTGCAAGAGACAGGCAACCACTGCTGCTG 1233

DB 828 TTCATGAGAGCATAGATGATGATGATGATGATGATGATGATGATGATGATGATG 887

QY 1234 CACTGCAAGATGGGCTCAGCGGCTCAGCGGCAAGTGTGCTTATGCTATGAGAG 1293

DB 888 CACTGCAAGATGGGCTCAGCGGCTCAGCGGCAAGTGTGCTTATGCTATGAGAG 947

QY 1294 TACGATGACGCTTGAGCAGAGCGGCTGCGGCAAGTGTGCTTATGCTATGAG 1353

DB 948 AAAGGCTGAGGCTGAGAGAGGCTTGAAGTGTGAAGAGGCGGACATCATCTCTG 1007

QY 1354 CCCAACCTGCTTCTGCGGCACTGCAAGTGTGCTTATGCTATGAGGCTGAGGCA 1408

DB 1008 CCCAACTTCACTGAGGAGGCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1062


```

RESULT 4
US-09-702-705-803
; Sequence 803, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 803
; LENGTH: 1238
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-803

Query Match 4.3%; Score 85.4; DB 4; Length 1238;
Best Local Similarity 50.4%; Pred. No. 3.1e-10;
Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGGAAACGACAACTGGAGAGCTGCAG 1053
DB 648 ATCCTTCCCTTCTTACCTGGCAGTGCTTACCATGCTGCCGAGAGACATGCTGAC 707
QY 1054 AGGAACAGGGTCAACCATCTGAACATGGCCCCGGAGATGACAACTTACCTGAG 1113
DB 708 GCCCTGGGCATCAGCGCTCTGTGAATGTCTCTCGACTGCCCAACCACTTTGAAGA 767
QY 1114 CGCTTACCTACCACAATGTGGCCTCTGGATGAGAGTGGCCCACTGCTGCCGAC 1173
DB 768 CACTATCAGTACAGTGCATCCAGTGAAGATAACCAAGGCCGACATCAGTCTCTGG 827
QY 1174 TGAAGAGAGCACCCTTCATTGAGGCTGCAAGAGCAGAGGCCACCCACGTGCTGTC 1233
DB 828 TTCATGGAAGCCATAGAGTACATCGATGCCGTGAAGACTGCCGTGGCGTGTGTG 887
QY 1234 CACTGCAAGATGGGCGTCAAGCGGCTCAGCGGCCACAGTGTGGCTATGCCATGAAGCAG 1293
DB 888 CACTGCCAGGCGGCATCTCGGGGTGGCCACCACATCTGCCCTGAGCTTACCTGATGATGAG 947
QY 1294 TACGAATGCAGCCTTGAGCAGAGGCCCTGCGCCACGTGACAGAGCTCCGGCCCATCGCCGC 1353
DB 948 AAACGGGTGAGGCTGAGAGAGGCTTCGAGTTGCTTAAGCAGGCCGACGATCATCTCG 1007
QY 1354 CCCAACCCCTGCTTCTCGCCGACAGCTGCAGATCTACCAAGGCGATCCTGACGGCCA 1408
DB 1008 CCCAATTCAAGCTTATGAGGGGACAGCTGCTGAGTTCCAGTCCAGGTGCTGCCA 1062

RESULT 5
US-09-736-457-803
; Sequence 803, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc

```

```

/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 803
/ LENGTH: 1238
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-09-736-457-803

Query Match          4.3%; Score 85.4; DB 4; Length 1238;
Best Local Similarity 50.4%; Pred. No. 3.1e-10;
Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAACGCAAACTTGAGGAGCTGCAG 1053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 648 ATCCTTCCCTTCCTTACCTCGCGCAGTGCCTACCATGTCTGCCGAGAGACATCTGAC 707
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1054 AGAACAGGGTACCCACATCTTGAACATGGCCCGAGATTGACAACTTCTACCTTGAG 1113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 708 GCCCTGGCATCAGCGCTCTGTGAATGTCTCTCGGACTGCCCAACCACTTTGAAGA 767
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1114 CGCTTACCTACCACATGTGCGCCTCTGGGATGAGAGTGGCCACGCTGTGCCGAC 1173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 768 CACTATCAGTACAGTGCATCCAGTGAAGATTAACCAAGCCGACATCAGCTCTGG 827
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1174 TGAAGAGAGCGCACCCCTTCATTGAGGCTGCAAGACACAGGGCACCCACGTGCTGTC 1233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 828 TTCAATGAAGCCATAGATACATCATGATGCCGTGAAGACTGCCGTGGCGCGTCTGTG 887
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1234 CACTGCAGATGGGCGCTCAGCCGCTCAGCGGCCACAGTGTGGCTATGCAAGACAG 1293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 888 CACTGCCAGGCGGGCATCTCGCGGTGGGCCACCACTCTGCTTACCTGATGATGAAG 947
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1294 TAGCAATGAGCCTGGAGCAGGCGCCTGCGCCACGTGACAGGAGCTCCGGCCCATGCCCC 1353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 948 AAACGGGTGAGGCTGAGGAGGCGCTTTCAGTTCTTTAAGCAGCGCCGACATCATCTCG 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1354 CCCAACCCCTGGCTTCTCTGCGCCAGCTGCAGATCTAACAGGGCATCTGACGGCCA 1408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1008 CCCAATTGACCTTCATGAGGGCAGCTGTGCACTTCAGTCCAGTCCAGGTGCTGGCCA 1062
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-09-702-705-825
/ Sequence 825, Application US/09702705
/ Patent No. 6504010
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C14
/ CURRENT APPLICATION NUMBER: US/09/702,705
/ CURRENT FILING DATE: 2000-10-30
/ NUMBER OF SEQ ID NOS: 1833
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 825
/ LENGTH: 2064
/ TYPE: DNA

```

ORGANISM: Homo sapiens
US-09-702-705-825

Query Match 4.3%; Score 85.4; DB 4; Length 2064;
Best Local Similarity 50.4%; Pred. No. 3.6e-10;
Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACTCTACCTGGGCTCAGAGTGAACGCAACCACTGAGAGAGCTGCAG 1053
DB 991 ATCCTTCCCTTCTCTACCTCGGCACTGCCCTACCATGTGCCCCGAGAGACATGCTGCAC 1050
QY 1054 AGGAACAGGGTCAACCCACATCTTGAACATGGCCCCGGAGATTGACAACTTCTAACCTGAG 1113
DB 1051 GCCCTGGGATCAAGGCTCTGTGTAATGTCTCTCGGACTGCCCAAACTTGAAGGA 1110
QY 1114 CGCTTCACTTACCAACAATGTGCGCTCTGGGATGAGAGTGGCCAGCTGCTGCCGAC 1173
DB 1111 CACTATCAGTACAAAGTGCATCCCACTGGAAGATAACCAAGGCCGACATGACTCTCGG 1170
QY 1174 TGAAGAGAGCGACCGCTTCAATGAGGCTGCAAGAGACAGGCAACCCAGCTGCTGTC 1233
DB 1171 TTCATGGAAGCCATAGAGTACATGATGCCGTGAAGAGACTGCCGTGGCGCTGCTGTG 1230
QY 1234 CACTGCAAGATGGGCGTCAAGCCGCTCAGCGGCCACAGTCTGCGCTATGCTGAAGCAG 1293
DB 1231 CACTGCCAGGGGGCATCTCGCGGTGGGCCACCATCTGCTGCTTACCTGATGATGAAG 1290
QY 1294 TACGAATGACGCTTGAGAGAGCGCTGCGCCACAGTGAAGAGCTCCGGCCATGCGCCGC 1353
DB 1291 AACGGGTGAGGCTGAGAGAGCGCTTGAAGTTCGTTAAGCAAGCGCCGACGATCATCTCG 1350
QY 1354 CCCAACCTGGCTTCTCTGCGCCAGCTGAGATCTACAGGCGCATCTGACGGCCA 1408
DB 1351 CCCACTTCAGCTTCATGAGGCGAGCTGCTGAGTTGAGTCCAGGTGCTGCCA 1405

RESULT 7

US-09-736-457-825

Sequence 825, Application US/09736457

Patent No. 6509448

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary

APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darrick

APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun

APPLICANT: Wang, Aijun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15

CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for windows Version 3.0

SEQ ID NO 825

LENGTH: 2064

TYPE: DNA

ORGANISM: Homo sapiens

US-09-736-457-825

Query Match 4.3%; Score 85.4; DB 4; Length 2064;
Best Local Similarity 50.4%; Pred. No. 3.6e-10;
Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACTCTACCTGGGCTCAGAGTGAACGCAACCACTGAGAGAGCTGCAG 1053
DB 991 ATCCTTCCCTTCTCTACCTCGGCACTGCCCTACCATGTGCCCCGAGAGACATGCTGCAC 1050
QY 1054 AGGAACAGGGTCAACCCACATCTTGAACATGGCCCCGGAGATTGACAACTTCTAACCTGAG 1113

DB 1051 GCCCTGGGATCAAGGCTCTGTGTAATGTCTCTCGGACTGCCCAAACTTGAAGGA 1110
QY 1114 CGCTTCACTTACCAACAATGTGCGCTCTGGGATGAGAGTGGCCAGCTGCTGCCGAC 1173
DB 1111 CACTATCAGTACAAAGTGCATCCCACTGGAAGATAACCAAGGCCGACATGACTCTCGG 1170
QY 1174 TGAAGAGAGCGACCGCTTCAATGAGGCTGCAAGAGACAGGGCAACCCAGCTGCTGTC 1233
DB 1171 TTCATGGAAGCCATAGAGTACATGATGCCGTGAAGAGACTGCCGTGGCGCTGCTGTG 1230
QY 1234 CACTGCAAGATGGGCGTCAAGCGCTCAGCGGCCACAGTCTGCGCTATGCTGAAGCAG 1293
DB 1231 CACTGCCAGGGGGCATCTCGCGGTGGGCCACCATCTGCTGCGCTTACCTGATGATGAAG 1290
QY 1294 TACGAATGACGCTTGAGAGAGCGCTGCGCCACAGTGAAGAGCTCCGGCCATGCGCCGC 1353
DB 1291 AACGGGTGAGGCTGAGAGAGCGCTTGAAGTTCGTTAAGCAGCGCCGACGATCATCTCG 1350
QY 1354 CCCAACCTGGCTTCTCTGCGCCAGCTGAGATCTACAGGCGCATCTGACGGCCA 1408
DB 1351 CCCACTTCAGCTTCATGAGGCGAGCTGCTGAGTTGAGTCCAGGTGCTGCCA 1405

RESULT 8

US-09-702-705-826

Sequence 826, Application US/09702705

Patent No. 6504010

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary

APPLICANT: Vedvick, Tom

APPLICANT: Carter, Darrick

APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C14

CURRENT APPLICATION NUMBER: US/09/702,705

CURRENT FILING DATE: 2000-10-30

NUMBER OF SEQ ID NOS: 1833

SOFTWARE: FastSeq for windows Version 3.0

SEQ ID NO 826

LENGTH: 2109

TYPE: DNA

ORGANISM: Homo sapiens

US-09-702-705-826

Query Match 4.3%; Score 85.4; DB 4; Length 2109;
Best Local Similarity 50.4%; Pred. No. 3.6e-10;
Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACTCTACCTGGGCTCAGAGTGAACGCAACCACTGAGAGAGCTGCAG 1053
DB 991 ATCCTTCCCTTCTCTACCTCGGCACTGCCCTACCATGTGCCCCGAGAGACATGCTGCAC 1010
QY 1054 AGGAACAGGGTCAACCCACATCTTGAACATGGCCCCGGAGATTGACAACTTCTAACCTGAG 1113
DB 1011 GCCCTGGGATCAAGGCTCTGTGTAATGTCTCTCGGACTGCCCAAACTTGAAGGA 1070
QY 1114 CGCTTCACTTACCAACAATGTGCGCTCTGGGATGAGAGTGGCCAGCTGCTGCCGAC 1173
DB 1071 CACTATCAGTACAAAGTGCATCCCACTGGAAGATAACCAAGGCCGACATGACTCTCGG 1130
QY 1174 TGAAGAGAGCGACCGCTTCAATGAGGCTGCAAGAGACAGGGCAACCCAGCTGCTGTC 1233
DB 1131 TTCATGGAAGCCATAGAGTACATGATGCCGTGAAGAGACTGCCGTGGCGCTGCTGTG 1190
QY 1234 CACTGCAAGATGGGCGTCAAGCGCTCAGCGGCCACAGTCTGCGCTATGCTGAAGCAG 1293

Db 1191 CACTGCCAGGGGGCATCTGCGGGTCGGCCACCACTGCTGGCCTACCTGATGTAAG 1250
QY 1294 TACGATGCAAGCCTGGAGCAGGCCCCCTGCGCCACGTGCAAGAGCTCCGGCCATCGCCCGC 1353
Db 1251 AAACGGGTGAGGCTGGAGAGGCCCCCTTGAGTTGTTAAGCAGCGCCGACATCATCTCG 1310
QY 1354 CCCAACCTTGCTTCTGCGCCGAGCTGCAAGATTAACAGGGCATCTGACGGCCA 1408
Db 1311 CCCAATTGACCTTCAATGGGGAGCTGCTGCAATTGAGTCCAGGTGCTGGCCA 1365

RESULT 9

US-09-736-457-826
; Sequence 826, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 826
; LENGTH: 2109
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-826

Query Match 4.3%; Score 85.4; DB 4; Length 2109;
Best Local Similarity 50.4%; Pred. No. 3.6e-10;
Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAACGACCAACCTGAGAGAGCTGCAG 1053
Db 951 ATCTTCCCTTCTCTACCTGGGCACTGCTACCACTGCTGCCGGAGAGACATGCTGAC 1010
QY 1054 AGGACAGGGTACCCACATCTTGAACATGGCCCCGGAGATTGACAACTTCTACCTGAG 1113
Db 1011 GCCCTGGGCATCAGCGCTGTGTAATGTCTCTCGGACTGCCCAACCACTTTGAAGGA 1070
QY 1114 CGCTTCACTTACCAATGTGGCTCTGGGATGAGAGTGGCCCCAGCTGCTGCCGAC 1173
Db 1071 CACTATCAGTACAAGTGCATGCCAGTGAAGATAACCAAGGCCGACATCAGCTCCTGG 1130
QY 1174 TGAAGAGAGCGCACCCTTCAATGAGGCTGCAAGAGCAGAGGACACCAAGCTGCTGTC 1233
Db 1131 TTCATGGAAGCCATAGAGTACATGATGCCGTGAAGACTGCCGTGGGCGCTGCTGTG 1190
QY 1234 CACTGCAAGATGGGCTGAGCGCTCAGCGGCCACAGTGTGGCTATGCCATGAAGCAG 1293
Db 1191 CACTGCAAGCGGGCATCTGGGCTGCGGCCACCATCTGCTGGCTACGTATGATGAAG 1250
QY 1294 TACGAATGACGCTGAGCAGGCCCCCTGCGCCACGTGCAAGAGCTCGGCCCATGCGCCG 1353
Db 1251 AAACGGGTGAGGCTGGAGAGGCCCCCTTGAGTTGTTAAGCAGCGCCGACATCATCTCG 1310
QY 1354 CCCAACCTTGCTTCTGCGCCGAGCTGCAAGATTAACAGGGCATCTGACGGCCA 1408
Db 1311 CCCAATTGACCTTCAATGGGGAGCTGCTGCAATTGAGTCCAGGTGCTGGCCA 1365

RESULT 10

US-09-016-434-1100
; Sequence 1100, Application US/09016434
; Patent No. 650938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Selhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 91255784
US-09-016-434-1100

Query Match 4.3%; Score 85.4; DB 4; Length 2240;
Best Local Similarity 50.4%; Pred. No. 3.7e-10;
Matches 209; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACCCTCTACCTGGGCTCAGAGTGAACGACCAACCTGAGAGAGCTGCAG 1053
Db 714 ATCTTCCCTTCTCTACCTGGGCACTGCTACCACTGCTGCCGGAGAGACATGCTGAC 773
QY 1054 AGGACAGGGTACCCACATCTTGAACATGGCCCCGGAGATTGACAACTTCTACCTGAG 1113
Db 774 GCCCTGGGCATCAGCGCTGTGTAATGTCTCTCGGACTGCCCAACCACTTTGAAGGA 833
QY 1114 CGCTTCACTTACCAATGTGGCTCTGGGATGAGAGTGGCCCCAGCTGCTGCCGAC 1173
Db 834 CACTATCAGTACAAGTGCATGCCAGTGAAGATAACCAAGGCCGACATCAGCTCCTGG 893
QY 1174 TGAAGAGAGCGCACCCTTCAATGAGGCTGCAAGAGCAGAGGACACCAAGCTGCTGTC 1233
Db 894 TTCATGGAAGCCATAGAGTACATGATGCCGTGAAGACTGCCGTGGGCGCTGCTGTG 953
QY 1234 CACTGCAAGATGGGCTGAGCGCTCAGCGGCCACAGTGTGGCTATGCCATGAAGCAG 1293
Db 954 CACTGCAAGCGGGCATCTGGGCTGCGGCCACCATCTGCTGGCTACGTATGATGAAG 1013
QY 1294 TACGAATGACGCTGAGCAGGCCCCCTGCGCCACGTGCAAGAGCTCGGCCCATGCGCCG 1353

Db 1014 AACGGGTGAGGCTGGAGAGGCGCTTCGAGTTCGTTAAGCAGCCCGCAGCATCTCG 1073
QY 1354 CCCAACCCCTGGCTTCCTGCGCCAGCTGCAGATCTACCGAGCATCTGACGCCA 1408
Db 1074 CCCAACTTCACTTTCATGGAGAGCTGCTGAGTTCGATCCAGGTGCTGGCCA 1128

RESULT 11

US-09-922-146-3
; Sequence 3, Application US/09922146
; Patent No. 656133
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowse
; APPLICANT: Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF DUAL SPECIFIC PHOSPHATASE 9 EXPRESSION
; FILE REFERENCE: RTS-0252
; CURRENT APPLICATION NUMBER: US/09/922,146
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 3
; LENGTH: 2303
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (114)...(1268)
US-09-922-146-3

Query Match 4.2%; Score 84; DB 4; Length 2303;
Best Local Similarity 51.9%; Pred. No. 7.7e-10;
Matches 217; Conservative 0; Mismatches 195; Indels 6; Gaps 1;

QY 990 CCGCATCTTCCCGCACTCTACCTGGGCTCAGAGTGAAGCAGCAAACTGAGAGCT 1049
Db 728 CCAGATCCCTGCCCACTCTATCTGGGAGTGCCCGGATTCGCCCAATTGAGAGCT 787
QY 1050 GCAGAGGAGAGGCTCACCACATTTGAACATGAGCCCGGAGATTGACAACTTCTACCC 1109
Db 788 GGCCAACTGGGCATCCGCTACATCTCAATGTACCCCAACTCCCAACTTCTTGA 847
QY 1110 TGAGCG-----CTTCACTACCAATGTGCGCTCTGGGATGAGAGTGGCCAGCT 1163
Db 848 GAAGATGTGATTTCACTACAGAGATCCCATCTCCGACCACTGGAGCGAAGACT 907
QY 1164 GCTGCGGCTGGAAGAGAGCGCAGCTTCTATGAGGCTGCAAGAGCAGGCGACCCA 1223
Db 908 GTGCGGCTTTTCCGAGGCCATTGAGTTCATTGATGAGGCTTGTCCAGACTGCGG 967
QY 1224 CGTCTGCTCACTGCAAGATGGGCTCAGCGCTCAGCGGCCACAGTGTGCTATGC 1283
Db 968 GGTGCTGCTCACTGCTTGGCGGGGTGAGCGGCTTGTCACTGCTGCTTACCT 1027
QY 1284 CATGAGCAGTGAAGATGAGAGCGGCTGAGAGCGGCTGCGCCACGTGAGAGAGCTCCGCC 1343
Db 1028 CATGAGAGCTCCACCTCTCTCTCAAGCATGCTTATGAGCTGCTCAAGAGAGAAAGTC 1087
QY 1344 CATGCGCGCGCCCAACCTGCTTCTCTGCGCCAGCTGCAAGATCTACAGGAGCTCTG 1401
Db 1088 TAACATCTCCCAACTTCACTTCACTGAGGCGAGTGTGACTTTGAGCGAGCTTG 1145

RESULT 12

US-09-702-705-801
; Sequence 801, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-801

Query Match 4.2%; Score 83.8; DB 4; Length 1619;
Best Local Similarity 50.1%; Pred. No. 7.7e-10;
Matches 208; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 994 ATCTTCCCACTCTTACCTGGGCTCAGAGTGAACGCAAACTGAGAGAGCTGCAG 1053
Db 320 ATCTTCCCTTCTCTTACCTGGGCTGCTTACCATGCTGCGCGAGAGACATGCTGAC 379
QY 1054 AGAAGAGGTCACCCACATCTTGAACATGCGCGGAGATTGACAACTTCTACCTGAG 1113
Db 380 GCCCTGGCATCAGGCTCTGTGATGTCTCTCGAGTGCACCAACCACTTGAAGA 439
QY 1114 CGCTTACCTACCAATGTGCGCTCTGGATGAGAGTGGCCAGCTGCTGCGCAC 1173
Db 440 CACTATCAGTACAGTGCATCCCGAGGAGATAACCAAGCCGACATGACTCTG 499
QY 1174 TGAAGAGAGCCACCGCTTCAATTGAGGCTGCAAGAGCAGGCAACCACTGCTGTC 1233
Db 500 TTCATGAGAGCATAGAGTACATGATGCGGTGAAGACTGCGGTGGCGCTGCTG 559
QY 1234 CACTGCAATGGCGCTCAGCGCTCAGCGGCGCCACAGTGTGCTATGCAAGAGAG 1293
Db 560 CACTGCCAGCGGCATCTCGCGGTGCGCCACCATCTGCTGCTACCTGATGAAG 619
QY 1294 TAGAATGACAGCTGAGAGAGCGCTGCGCCACAGTGCAGAGAGCTCCGCGCATGCGCCG 1353
Db 620 AACGGGTGAGGCTGAGAGAGCGCTTTCAGATTGTTAAGCAGCGCGCAGCATTAATCTCG 679
QY 1354 CCCAACCTGCTTCTGCGCCAGCTGCAAGATCTACAGGAGCATCTGACGCCA 1408
Db 680 CCCAACTTCACTTCACTGAGGCGAGCTGCTGAGTTCAGTCCAGGTGCTGCGCA 734

RESULT 13

US-09-736-457-801
; Sequence 801, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 801
; LENGTH: 1619
; TYPE: DNA

ORGANISM: Homo sapien
US-09-736-457-801

Query Match 4.2%; Score 83.8; DB 4; Length 1619;
Best Local Similarity 50.1%; Pred. No. 7.7e-10;
Matches 208; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACCCTTACCTTGGGCTCAGAGTGAACGCAAAACCTGAGAGCTGCAG 1053
Db 320 ATCCTTCCCTTCTTACCTCGGAGTGCCTACCATGCTGCCCGAGAGACATGCTGCAC 379
QY 1054 AGGAACAGGGTCAACCAATCTTGAACATGGCCCGGAGATTGAACAATTCTACCTTGAG 1113
Db 380 GCCCTGGGCATCAGGGCTCTGTGAATGTCCTCGGACTGCCCAAAACCACTTTGAAGGA 439
QY 1114 CGCTTACCTACCAATATGTGCGCCTCTGGAGTGAAGAGTCCGCCCAAGCTGTGCTGCCGAC 1173
Db 440 CACTATCAGTACAGAGTGCATCCAGTGAAGATAACCAAGGCGGACATCAGCTCCTGG 499
QY 1174 TGAAGAGAGACCGCAGCCTTTCATTTAGAGCTGCAAGAGACACAGGCAACCCAGCTGTGTC 1233
Db 500 TTCATGGAAGCCATAGATACATGATGATCCGTGAAGACTGCCGTGGCGCTGTGTTG 559
QY 1234 CACTGCAAGATGGCGCTCAGCCGCTCAGCGGCCACAGTGTCTGCTATGCCATGAAGCAG 1293
Db 560 CACTGCCAGGCGGGCATCTCGCGGTGCGCCACCATCTGCTGCTGCTACCTGATGATGAAG 619
QY 1294 TAGCAATGACGCTTGAGAGCAGGCGCTGCGGCCACGTGACAGAGTCCGGCCCATCGCCGC 1353
Db 620 AAACGGGTGAGGCTGAGAGAGGCGCTTGAAGTCTGTTAAGACAGGCGCCGACATTAATCTCG 679
QY 1354 CCCAACCCCTGCTTCTCTGCGCCAGCTGAGATCTACCAGGCAATCTGACGGCCA 1408
Db 680 CCCAACTTCACTTTCATGGGGCAGCTGTGCACTTCCAGTCCAGTGTGCGCA 734

RESULT 14

US-09-702-705-804
Sequence 804, Application US/09702705
Patent No. 6504010
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedwick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C14
CURRENT APPLICATION NUMBER: US/09/702,705
CURRENT FILING DATE: 2000-10-30
NUMBER OF SEQ ID NOS: 1833
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 804
LENGTH: 4637
TYPE: DNA
ORGANISM: Homo sapiens
US-09-702-705-804

Query Match 4.2%; Score 83.8; DB 4; Length 4637;
Best Local Similarity 50.1%; Pred. No. 1e-09;
Matches 208; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACCCTTACCTTGGGCTCAGAGTGAACGCAAAACCTGAGAGCTGCAG 1053
Db 320 ATCCTTCCCTTCTTACCTCGGAGTGCCTACCATGCTGCCCGAGAGACATGCTGCAG 379
QY 1054 AGGAACAGGGTCAACCAATCTTGAACATGGCCCGGAGATTGAACAATTCTACCTTGAG 1113

Db 380 GCCCTGGGCATCAGGGCTCTGTGAATGTCCTCTCGGACTGCCCAAAACCACTTTGAAGGA 439

QY 1114 CGCTTACCTTACCACAAATGTGCGCCTCTGGAGTGAAGAGTGGGCCCAAGCTGTGCCGAC 1173
Db 440 CACTATCAGTACAGAGTGCATTCGCCAGTGAAGATTAACCAAGGCGGACATCAGCTCCTGG 499
QY 1174 TGAAGAGAGCAGCCGCTTTCATTTAGAGCTGCAAGAGACACAGGCAACCCAGCTGTGTC 1233
Db 500 TTCATGGAAGCCATAGATACATGATGATCCGTGAAGACTGCCGTGGCGCTGTGTTG 559
QY 1234 CACTGCAAGATGGCGCTCAGCCGCTCAGCGGCCACAGTGTGCTGCTATGCCATGAAGCAG 1293
Db 560 CACTGCCAGGCGGGCATCTCGCGGTGCGCCACCATCTGCTGCTGCTACCTGATGATGAAG 619
QY 1294 TAGCAATGACGCTTGAGAGAGGCGCTGCGGCCAGTGAAGAGTCCGGCCCATCGCCGC 1353
Db 620 AAACGGGTGAGGCTGAGAGAGGCGCTTGAAGTCTGTTAAGACAGGCGCGGACATTAATCTCG 679
QY 1354 CCCAACCCCTGCTTCTCTGCGCCAGCTGACATCTACCAGGCAATCTGACGGCCA 1408
Db 680 CCCAACTTCACTTTCATGGGGCAGCTGTGCACTTCCAGTCCAGTGTGCGCA 734

RESULT 15

US-09-736-457-804
Sequence 804, Application US/09736457
Patent No. 6509448
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedwick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
APPLICANT: Wang, Aijun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.478C15
CURRENT APPLICATION NUMBER: US/09/736,457
CURRENT FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 1864
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 804
LENGTH: 4637
TYPE: DNA
ORGANISM: Homo sapiens
US-09-736-457-804

Query Match 4.2%; Score 83.8; DB 4; Length 4637;
Best Local Similarity 50.1%; Pred. No. 1e-09;
Matches 208; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 994 ATCTTCCCCCACCCTTACCTTGGGCTCAGAGTGAACGCAAAACCTGAGAGCTGCAG 1053
Db 320 ATCCTTCCCTTCTTACCTCGGAGTGCCTACCATGCTGCCCGAGAGACATGCTGCAG 379
QY 1054 AGGAACAGGGTCAACCAATCTTGAACATGGCCCGGAGATTGAACAATTCTACCTTGAG 1113
Db 380 GCCCTGGGCATCAGGGCTCTGTGAATGTCCTCTCGGACTGCCCAAAACCACTTTGAAGGA 439
QY 1114 CGCTTACCTTACCACAAATGTGCGCCTCTGGAGTGAAGAGTCCGCCCAAGCTGTGCCGAC 1173
Db 440 CACTATCAGTACAGAGTGCATCCAGTGAAGATAACCAAGGCGGACATCAGCTCCTGG 499
QY 1174 TGAAGAGAGCAGCCGCTTTCATTTAGAGCTGCAAGAGACACAGGCAACCCAGCTGTGTC 1233
Db 500 TTCATGGAAGCCATAGATACATGATGATCCGTGAAGACTGCCGTGGCGCTGTGTTG 559
QY 1234 CACTGCAAGATGGCGCTCAGCCGCTCAGCGGCCACAGTGTGCTATGCCATGAAGCAG 1293

DB 560 CACTGCCAGCGCGGCACTCTCGGGTCCGCAACCATCTGCTGGCCTACCTGATGATGAG 619
QY 1294 TACGAATGCAGCCTTGAGCAGGCCCTTGCGCAAGTGACAGAGCTCCGGCCATCGCCCGC 1353
DB 620 AACGGGTGAGGCTTGAGAGAGCCCTTCGAGTTGTTAAGCAGCGCCGACATTATCTCG 679
QY 1354 CCCAACCTTGCTTCTGCGCCAGCTGCAGATCTACAGGGGCATCTGACGGCCA 1408
DB 680 CCCAAGTTCAAGCTTATGAGGCGAGCTGCTGAGTTGAGTCCAGGTGCTGGCCA 734

Search completed: January 15, 2004, 18:31:49
Job time : 146 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 06:48:01 ; Search time 21 Seconds
(without alignments)
3017.865 Million cell updates/sec

Title: US-09-955-732-2
Perfect score: 3412
Sequence: 1 MALVTVSRSPPGSGASTPVG.....RRKKVVRQASVHDSGEEGEA 659
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: pirl:*
2: pirl2:*
3: pirl3:*
4: pirl4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	269.5	7.9	220	JC7885	low-molecular-mass
2	224	6.6	384	I38890	dual specificity p
3	214	6.3	226	T21380	hypothetical prote
4	210.5	6.2	272	T18915	hypothetical prote
5	210.5	6.2	619	T15969	hypothetical prote
6	209.5	6.1	393	A56947	dual specificity p
7	209	6.1	314	B57126	dual specificity p
8	206.5	6.1	394	A56115	dual specificity p
9	204.5	6.0	303	T46405	hypothetical prote
10	202	5.9	314	A57126	dual specificity p
11	202	5.9	367	S29090	dual specificity p
12	196.5	5.8	223	T49365	protein tyrosine p
13	194.5	5.7	365	T32494	hypothetical prote
14	194	5.7	367	S24411	dual specificity p
15	194	5.7	367	S52265	dual specificity p
16	193	5.7	489	S58725	dual specificity p
17	188.5	5.5	205	T49364	protein tyrosine p
18	187	5.5	330	T39698	protein tyrosine p
19	186.5	5.5	185	A47196	dual specificity p
20	186.5	5.5	286	F88481	protein C16A3.1 [i
21	185	5.4	283	G84458	probable protein p
22	179	5.2	142	T03074	dual specificity p
23	177.5	5.2	276	T48906	protein-tyrosine-p
24	176	5.2	2472	E83594	still frameshift p
25	175.5	5.1	364	S31304	protein-tyrosine-p
26	169	5.0	186	T16056	hypothetical prote
27	155.5	4.6	4957	T03455	ALR protein - huma
28	155.5	4.6	5262	T03454	ALR protein - huma
29	154	4.5	292	S41012	hypothetical prote

30	153	4.5	580	2	T18439	hypothetical prote
31	153	4.5	600	2	T18446	hypothetical prote
32	143	4.2	771	2	T47666	phosphatase-like p
33	141	4.1	3938	2	T42761	Basoon protein -
34	140.5	4.1	1805	2	A34736	nestin - rat
35	135.5	4.0	1571	2	T00062	hypothetical prote
36	135.5	4.0	1606	2	T34073	paranemin - chicke
37	133	3.9	1346	2	A57376	probable regulator
38	132.5	3.9	837	2	T02761	outer arm dynein i
39	131.5	3.9	807	1	S44538	probable protein-t
40	131.5	3.9	2175	1	S03170	homeotic protein c
41	130	3.8	990	2	T14756	hypothetical prote
42	130	3.8	1426	2	T00337	hypothetical prote
43	128.5	3.8	1186	2	G69708	chromosome segrega
44	128	3.8	209	1	S48459	probable dual spec
45	128	3.8	1366	2	B86292	F7H2.12 protein -

ALIGNMENTS

RESULT 1
JC7885 low-molecular-mass dual-specificity phosphatase-2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 09-Dec-2002 #sequence_revision 09-Dec-2002 #text_change 31-Mar-2003
C;Accession: JC7885
R;Nakamura, K.; Tanoue, K.; Satoh, T.; Takekawa, M.; Watanabe, M.; Shima, H.; Kikuchi, K
J. Biochem. 132, 463-470, 2002
A;Title: A novel low-molecular-mass dual-specificity phosphatase, LDP-2, with a natural
A;Reference number: JC7885; MUID:22194259; PMID:12204117
A;Accession: JC7885
A;Molecule type: mRNA
A;Residues: 1-220 <NAK>
A;Cross-references: DDBJ:AB038769; DDBJ:AB038770
C;Comment: This enzyme, a novel member of the low-molecular-mass dual-specificity phosph
ein kinase signaling.
C;Genetics:
A;Gene: ldp-2

Query Match 7.9%; Score 269.5; DB 2; Length 220;
Best Local Similarity 38.0%; Pred. No. 2.9e-10;
Matches 62; Conservative 30; Mismatches 58; Indels 13; Gaps 2;

QY 332 IFPHLYGSEWNAANLEIQRNRVTHILNMAREIDNPFYPERFTYHNVRLMDESAQLLPH 391
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 68 IKPWLLIGSQDAAHDLRLKHKVTHILNVAYGVENAFLEFYTITISLDVPEITNLSY 127
QY 392 WKETHRTFBAARAQGTHTVLVCKMGVSRSAATVLAAMKQYECSLQALRHVQELRPIAR 451
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 128 FPECFEFTFQAKLKDGVVAVHCNAGVSRAAAIVIGFLMSSEATFTTALSIVKEARPSIC 187
QY 452 FNPGLRQLQIYQGITLTSRQSHWVEQKVGVSPBEHPAPEVS 494
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 188 FNPGLRQLRTYQ-----VGKESNGG--DKVPAEDTT 217

RESULT 2
I38890 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 5 - human
N;Alternate names: dual specificity phosphatase B23; dual-specificity phosphatase hvh-3;
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 11-Jun-1999
C;Accession: I38890; A55313
R;Kwak, S.P.; Dixon, J.E.
J. Biol. Chem. 270, 1156-1160, 1995
A;Title: Multiple dual specificity protein tyrosine phosphatases are expressed and regul
A;Reference number: A55432; MUID:95138103; PMID:7836374.
A;Accession: I38890
A;Molecule type: mRNA
A;Residues: 1-384 <RES>
A;Cross-references: EMBL:U16996; NID:g642012; PIDN:AAB06261.1; PID:g642013
A;Experimental source: placenta

C/Genetics:

A:Gene: CESP:F08B1.1

A:Introns: 22/2; 57/1; 99/1; 160/3; 252/2; 549/3

Query Match 6.2%; Score 210.5; DB 2; Length 619;

Best Local Similarity 20.9%; Pred. No. 7.6e-06;

Matches 113; Conservative 66; Mismatches 198; Indels 163; Gaps 18;

QY 268 QEQMEQAIRALMKVLD-----VSDLESVTSKEIRQALERL-----GLPL----- 308

Db 40 RERLEDTDKFRSVMTLEEQDPVSSLPTFPAPKRGKXLQTLTSSPINSSSPISSSSPTNG 99

QY 309 -----QYRDFID-----NOMLLVAQRDRASRIFPHLYGSEWMAANLEEL 350

Db 100 GFKQFAQQYPOLCESSEGMTRLPQSLSQPCLSQPTGDTLITPNLYGSGQIDSLDETML 159

QY 351 QGNRVTHILNMA-----REIDNFPERTYHNVRLWDEESAQLLPHWKETHRTIEA 401

Db 160 DALDISVIVINLSMTCPKSVCIKEDKNFM-----RIPVNDSYQEKLSPEFPMAYEFLEK 212

QY 402 ARAQGTHTLVHCKMGVSRSAATVLAYAMKQYECGLEQALRHVOELRPPIARPNGFLROLQ 461

Db 213 CRAGKKCLIHCLAGISRSPTLAISYIMRYMKGSDDAYRYVKERRPSISPNFMGQL 272

QY 462 IYQGIL-----TASR-----QSHWEQKV-----GCVSPEEHPA-- 490

Db 273 EVENVLIKDHLVDYNQASRPHRMDYGFSDLCPPKVPKSASNSCVFPGSTHDESSPSSP 332

QY 491 -----PEVST-----PFPP-----LPEE 503

Db 333 SVSEGSAASEPETSSSAASSSSTASAPPSMPTSEOGTSSGTAVNNGKRNMTMDIGLPHR 392

QY 504 PEGGEEKVVGMEESQ-AAPKEE-----PGPR-----PRINLRGVN 538

Db 393 PKALGLPSRIGTSVAELPSPSTELSLSFNGPEALAPSTPILNFTNCPFNSPIIIPVASSS 452

QY 539 RSISLLEPSLELESTSETSDMPEV-FSSHESHEEPL--QPFPOLARTKGQGVDRGQ 594

Db 453 REVILLTPTPAASSSSSTSEPSFDFSSFESSSSSSSIVENPFPASTEVPAAGSSSISIPSS 512

QY 595 PALKSRQSVTLQGSAYVANRTOAFQEQEQGGQGGEPCTISSTPRFRKVVROASVHDSG 654

Db 513 ---GSOSTPASASSSSAASRCRMKGFVKVFSKKAAPASTSTPASSTPGTSRAARPECLRSSG 569

RESULT 6

A56947 dual specificity phosphatase (EC 3.1.3.-) HVH2 - rat

N/Alternate names: mitogen-activated protein kinase phosphatase 2

C/Species: Rattus norvegicus (Norway rat)

C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-1999

C/Accession: A56947

R/Mistra-Press, A.; Rim, C.S.; Yao, H.; Roberson, M.S.; Stork, P.J.S.

J. Biol. Chem. 270, 14587-14596, 1995

A/Title: A novel mitogen-activated protein kinase phosphatase. Structure, expression, an

A/Reference number: A56947; MUID:95301550; PMID:7782322

A/Accession: A56947

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-393 <MIS>

C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C/Keywords: phosphoprotein; phosphoric monoester hydrolase

F/202-333/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

QY 392 WKETHRTIEARAQGTHTLVHCKMGVSRSAATVLAYAMKQYECGLEQALRHVOELRPIAR 451

Db 258 FMEAIEYIDAVKDCRGRVLVHCQAGISRSATICLAYLMKKRVRLLEAFEFVKQRSIIS 317

QY 452 PNPGLRQ-LQIYQGITPASRQSHWEQKVGVSPDEHPAPEVSTPFPPPL 500

Db 318 PNFSFMGQLQFESQVLTTSS-----CAAEASPSGFL 349

RESULT 7

B57126 dual specificity phosphatase (EC 3.1.3.-) 2 - mouse

N/Alternate names: mitogen-induced nuclear protein-tyrosine-phosphatase PAC-1

C/Species: Mus musculus (house mouse)

C/Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 24-Apr-1998

C/Accession: B57126

R/Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzsch, H.; Siebenlist, U.; Keli

Science 259, 1763-1766, 1993

A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.

A/Reference number: A57126; MUID:93206122; PMID:7681221

A/Accession: B57126

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-314 <ROH>

A/Cross-references: GB:L11330

C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C/Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase

F/180-311/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F/257/Active site: Cys (phosphocysteine intermediate) #status predicted

F/263/Binding site: substrate phosphate (Arg) #status predicted

Query Match

Best Local Similarity 6.1%; Score 209; DB 2; Length 314;

Matches 43; Conservative 33; Mismatches 53; Indels 0; Gaps 0;

QY 332 IFPHLYGSEWMAANLEELQNRVTHILNMAREIDNFPERTYHNVRIMDEESAQLLPH 391

Db 176 ILPFLYIGSCNHSDDLQGLQACGITTAVLNVSASCPNHFEGLFHYKSIPEVDNQWEISAW 235

QY 392 WKETHRTIEARAQGTHTLVHCKMGVSRSAATVLAYAMKQYECGLEQALRHVOELRPIAR 451

Db 236 FOEATISFIDSVKNSGGRVLVHCQAGISRSATICLAYLIQSHRVRLDEARDFVKQRRGVIS 295

QY 452 PNPGLRQQL 460

Db 296 PNFSFMGQL 304

RESULT 8

A56115 dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 4 - human

N/Alternate names: dual specificity phosphatase HVH2

C/Species: Homo sapiens (man)

C/Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 24-Apr-1998

C/Accession: A56115

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-394 <GUA>

A/Cross-references: GB:U21108

C/Genetics:

A/Gene: GDB:DUSP4; HVH2; MKP-2

A/Cross-references: GDB:433893

A/Map position: 8p21-8p11.2

C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity

C/Keywords: phosphoprotein; phosphoric monoester hydrolase

F/203-334/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>

F/260/Active site: Cys (phosphocysteine intermediate) #status predicted

F:286/Binding site: substrate phosphate (Arg) #status predicted

Query Match 6.1%; Score 206.5; DB 2; Length 394;
Best Local Similarity 23.2%; Pred. No. 7.3e-06;
Matches 82; Conservative 46; Mismatches 124; Indels 101; Gaps 7;

```
QY 222 GGSALTWASHYQERLNSEQCLNWTAMADLESIRPPSAEPGG----- 264
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 25 GGGAGGTGSHGTLGLPSGKCL-----LLDCRPFLAHSAGYILGSVNVRCNTIVRR 76

QY 265 -----SSEQ-----EQMEQATRAELMKVLVDVSDLESVTSKEIRQALEIRLGLPLQYRDF 314
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 77 AKGSVSELEQILPAEEVRRARLSGLYSAVIVYDERSPRAESLRE--DSTVSLVQALRRN 134

QY 315 ID-NQMLLVQDRAS----- 330
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 135 AERTDCLIKGYERFSSEYPERFCSKTALAIPPPVPSPATEPLDLCSSCGTPLHDQG 194

QY 331 ---RIFPHLYLGSEWNAANLEELQNRVTHILNMAREIDNFFPERFYHNVRLWDEESAQ 387
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 195 GPVEILPFLYLGSAVYHARRMDLALGITALLNVSSDCPNHFEGHYQKCIPIVEDHKKAD 254

QY 388 LLPHWKETHRFIEARAQGTHTLVHCKMGVSRSAATVLAAYAMKQYECSEQALRHVQELR 447
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 255 ISSWFMEALIEYIDAVKDCRGRLVHCOAGISRSATICLAYLMKKRVRLBEAFEFVKQRR 314

QY 448 FIARPNPGLRQLQIYQGITLTSRQSHVWEQKVGVSPEEHPAPEVSTPPPL 500
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 315 SIISPNFSFMQLLQFESQVLAT-----SCAAEAASPSGPL 350
```

RESULT 9

hypothetical protein DKFZp43401321.1 - human

C/Species: Homo sapiens (man)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 21-Jun-2002
C/Accession: T46405
R/Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A/Reference number: Z23034
A/Accession: T46405
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-303 <AAA>
A/Cross-references: EMBL:AL137704
A/Experimental source: adult testis; clone DKFZp43401321
C/Genetics:
A/Note: DKFZp43401321.1

Query Match 6.0%; Score 204.5; DB 2; Length 303;
Best Local Similarity 30.2%; Pred. No. 6.8e-06;
Matches 51; Conservative 30; Mismatches 71; Indels 17; Gaps 1;

```
QY 332 IFPHLYLGSEWNAANLEELQNRVTHILNMAREIDNFFPERFYHNVRLWDEESAQLLPH 391
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 108 ILPFLYLGSAVYHARRMDLALGITALLNVSSDCPNHFEGHYQKCIPIVEDNHKADISSW 167

QY 392 WKETHRFIEARAQGTHTLVHCKMGVSRSAATVLAAYAMKQYECSEQALRHVQELRPIAR 451
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 168 FMEALIEYIDAVKDCRGRLVHCOAGISRSATICLAYLMKKRVRLBEAFEFVKQRRSIIS 227

QY 452 PNPGLRQLQIYQGITLTSRQSHVWEQKVGVSPEEHPAPEVSTPPPL 500
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 228 PNFSEFMQLLQFESQVLAT-----SCAAEAASPSGPL 259
```

RESULT 10

A57126

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 2 - human
N/Alternate names: mitogen-induced nuclear tyrosine phosphatase; phosphatase of activated
C/Species: Homo sapiens (man)
C/Date: 03-Nov-1995 #sequence_revision 02-Jul-1996 #text_change 11-Jun-1999
C/Accession: A57126

R/Rohan, P.J.; Davis, P.; Moskaluk, C.A.; Kearns, M.; Krutzsch, H.; Siebenlist, U.; Kell
Science 259, 1763-1766, 1993

A/Title: PAC-1: a mitogen-induced nuclear protein tyrosine phosphatase.
A/Reference number: A57126; MUID:93206122; PMID:7681221
A/Accession: A57126
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-314 <ROH>
A/Cross-references: GB:L11329; NID:9559539; PIDN:AAA50779.1; PID:9292376
C/Genetics:
A/Gene: GDB:DUSP2
A/Cross-references: GDB:139200
A/Map position: 2q11-2q11
C/Function:
A/Description: catalyzes the hydrolysis of peptidyl-phosphoserine, -phosphothreonine, an
C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C/Keywords: nucleus; phosphoprotein; phosphoric monoester hydrolase
F:140-160/Region: nuclear location signal
F:180-311/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F:257/Active site: Cys (phosphocysteine intermediate) #status predicted
F:263/Binding site: substrate phosphate (Arg) #status predicted

Query Match 5.9%; Score 202; DB 1; Length 314;
Best Local Similarity 33.3%; Pred. No. 1e-05;
Matches 43; Conservative 31; Mismatches 55; Indels 0; Gaps 0;

```
QY 332 IFPHLYLGSEWNAANLEELQNRVTHILNMAREIDNFFPERFYHNVRLWDEESAQLLPH 391
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 176 ILPFLYLGSHSSDQLQACGITAIVNVASCPNHFEGLFYKSIPIVEDNQWEISAW 235

QY 392 WKETHRFIEARAQGTHTLVHCKMGVSRSAATVLAAYAMKQYECSEQALRHVQELRPIAR 451
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 236 FQEAIGFIDWVKNSGRVLVHCOAGISRSATICLAYLMQSRVRRLDEAFDFVKQRRGVIS 295

QY 452 PNPGLRQL 460
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 296 PNFSEFMQL 304
```

RESULT 11

S29090

dual specificity phosphoprotein phosphatase (EC 3.1.3.-) 1 - human
N/Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase, non
C/Species: Homo sapiens (man)
C/Date: 25-Feb-1994 #sequence_revision 02-May-1994 #text_change 11-Jun-1999
C/Accession: S29090; A53052
R/Keyse, S.M.; Emslie, E.A.
Nature 359, 644-647, 1992
A/Title: Oxidative stress and heat shock induce a human gene encoding a protein-tyrosine
A/Reference number: S29090; MUID:93024952; PMID:1406996
A/Accession: S29090
A/Molecule type: mRNA
A/Residues: 1-367 <KEY>

A/Cross-references: EMBL:X68277; NID:G29980; PIDN:CAA48338.1; PID:G29981
R/Kwak, S.P.; Hakes, D.J.; Martell, K.J.; Dixon, J.E.
J. Biol. Chem. 269, 3596-3604, 1994
A/Title: Isolation and characterization of a human dual specificity protein-tyrosine pho
A/Reference number: A53052; MUID:94148864; PMID:8106404
A/Accession: A53052
A/Molecule type: DNA
A/Residues: 1-367 <KWA>

A/Experimental source: leukocyte
A/Note: sequence extracted from NCBI backbone (NCBIN:143800, NCBIN:143802, NCBIN:143804
C/Genetics:
A/Gene: GDB:DUSP1; PTPN10
A/Cross-references: GDB:136197; OMIM:600714
A/Map position: 5q34-5q34
C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C/Keywords: heat shock; phosphoprotein; phosphoric monoester hydrolase; stress-induced
F:181-312/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F:256/Active site: Cys (phosphocysteine intermediate) #status predicted
F:264/Binding site: substrate phosphate (Arg) #status predicted

DB 297 PNFSEMGQLLOFESQVLA 314

RESULT 15

S52265
dual specificity phosphatase (EC 3.1.3.-) 1 - rat
N/Alternate names: protein-tyrosine-phosphatase CL100; protein-tyrosine-phosphatase, non
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 22-Jun-1999
C/Accession: S52265
R/Muda, M.; Schlegel, W.; Arkinstall, S.
submitted to the EMBL Data Library, January 1995
A/Description: Pathways regulating CL100 gene expression in pituitary cells.
A/Reference number: S52265
A/Accession: S52265
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-367 <MUD>
A/Cross-references: EMBL:X84004; NID:g642264; PIDN:CAA58828.1; PID:g642265
C/Superfamily: dual specificity phosphoprotein phosphatase 1; VH1-type dual specificity
C/Keywords: phosphoprotein; phosphoric monoester hydrolase
F/181-312/Domain: VH1-type dual specificity phosphoprotein phosphatase homology <VH1>
F/258/Active site: Cys (phosphocysteine intermediate) #status predicted
F/264/Binding site: substrate phosphate (Arg) #status predicted

Query Match 5.7%; Score 194; DB 2; length 367;
Best Local Similarity 30.4%; Pred. No. 4.2e-05;
Matches 42; Conservative 30; Mismatches 66; Indels 0; Gaps 0;

QY	332	IFPHLYLGSEWNAANLELQNRVTHILNMAREIDNFYPERFTYHNRLWDESAQLLP	391
DB	177	ILSFYLGSAVHASRKMIDALGITALINVSANCPNHFEHYQKSIPEVDNKHADISSW	236
QY	392	WKETHRFIEARAQGTHTLVHCKMGVSRSAATVLAYAMKOYECSLBOALRHVQELRPIAR	451
DB	237	FNEAIDFIDSIKIDAGGRVFVHCQAGISRSATICLAYIMRTNRVKLDEAFEFYKQRRSITS	296
QY	452	PNPGFLRQLQIYQGITLA	469
DB	297	PNFSFMGQLLOFESQVLA	314

Search completed: January 15, 2004, 06:50:27
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 06:41:46 ; Search time 18 Seconds
(without alignments)
1721.700 Million cell updates/sec

Title: US-09-955-732-2
Perfect score: 3412
Sequence: 1 MALVTVSRSPPGSGASTPVG.....RRKVVRAQSVHDSGEEGEA 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	7.6	625	DUS8_HUMAN	Q13202 homo sapien
2	254	7.4	217	DUSJ_HUMAN	Q8WTR2 homo sapien
3	253.5	7.4	665	DUSG_HUMAN	Q9BY84 homo sapien
4	252	7.4	663	DUS8_MOUSE	Q09112 mus musculu
5	239.5	7.0	482	DUSA_HUMAN	Q9Y6W6 homo sapien
6	233	6.8	384	DUS5_RAT	Q54838 rattus norv
7	232.5	6.8	483	DUSA_MOUSE	Q9ES80 mus musculu
8	229	6.7	295	DUSF_HUMAN	Q9H1R2 homo sapien
9	226.5	6.6	320	DUS7_MOUSE	Q91Z46 mus musculu
10	225.5	6.6	384	DUS9_HUMAN	Q99956 homo sapien
11	224	6.6	384	DUS5_HUMAN	Q16690 homo sapien
12	223.5	6.6	280	DUS7_RAT	Q63340 rattus norv
13	217.5	6.4	320	DUS7_HUMAN	Q16829 homo sapien
14	215.5	6.3	395	DUS4_RAT	Q62767 rattus norv
15	210.5	6.2	619	VHPI_CAEEL	Q10038 caenorhabdi
16	209	6.1	318	DUS2_MOUSE	Q05922 mus musculu
17	205.5	6.0	394	DUS4_HUMAN	Q13115 homo sapien
18	204	6.0	375	DUS4_CHICK	Q9PW71 gallus gall
19	202	5.9	314	DUS2_HUMAN	Q05923 homo sapien
20	202	5.9	367	DUS1_HUMAN	P28562 homo sapien
21	198.5	5.8	223	STYX_HUMAN	Q8WUJ0 homo sapien
22	196.5	5.8	223	STYX_MOUSE	Q60969 mus musculu
23	194	5.7	367	DUS1_MOUSE	P28563 mus musculu
24	194	5.7	367	DUS1_RAT	Q64653 rattus norv
25	193	5.7	489	MSGS_YEAST	P38590 saccharomyc
26	192.5	5.6	381	DUS6_MOUSE	Q9DBH1 mus musculu
27	191.5	5.6	340	DUSC_HUMAN	Q9UN16 homo sapien
28	191	5.6	188	DUS1_HUMAN	Q8NEJ0 homo sapien
29	191	5.6	198	DUSD_HUMAN	Q9U1I6 homo sapien
30	189.5	5.6	198	DUSD_MOUSE	Q9GYJ7 mus musculu
31	189.5	5.6	381	DUS6_RAT	Q64346 rattus norv
32	186.5	5.5	185	DUS3_HUMAN	P51452 homo sapien
33	186.5	5.5	381	DUS6_HUMAN	Q16828 homo sapien

34	184.5	5.4	198	DUSE_MOUSE	Q9J1Y7 mus musculu
35	182.5	5.3	198	DUSE_HUMAN	Q95147 homo sapien
36	181.5	5.3	125	DUSF_MOUSE	Q84V22 mus musculu
37	179.5	5.3	185	DUS3_MOUSE	Q967X3 mus musculu
38	177.5	5.2	276	PTP3_CHLEU	Q39491 chlamydomon
39	177.5	5.2	339	DUSC_MOUSE	Q960T2 mus musculu
40	175.5	5.1	364	PVH1_YEAST	Q02256 saccharomyc
41	154	4.5	292	Y042_CAEEL	P34680 caenorhabdi
42	147	4.3	912	A36D_DROME	Q9V3R1 drosophila
43	140.5	4.1	1805	NEST_RAT	P21263 rattus norv
44	131.5	3.9	598	CCBI_BOVIN	Q9MZ17 bos taurus
45	131.5	3.9	807	YB9T_YEAST	P38148 saccharomyc

ALIGNMENTS

RESULT 1
ID DUS8_HUMAN STANDARD; PRT; 625 AA.
AC Q13202;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase hvh-5).
GN DUSP8 OR VH5.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96009533; PubMed=7561881;
RA Martell K.J., Seasholtz A.F., Kwak S.P., Clemens K.K., Dixon J.E.;
RT "hvh-5: a protein tyrosine phosphatase abundant in brain that
inactivates mitogen-activated protein kinase.";
RT J. Neurochem. 65:1823-1833(1995).
RL J. Neurochem. 65:1823-1833(1995).
CC -!- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = a protein +
tyrosine + phosphate.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
phosphate.
CC -!- TISSUE SPECIFICITY: ABUNDANT IN BRAIN, HEART, AND SKELETAL MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -!- SIMILARITY: Contains 1 rhodanese domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U27193; AAA83151.1; -
DR HSSP; Q16828; 1MKP.
DR Genew; HGNC:3074; DUSP8.
DR MIM; 602038; -
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO; GO:0000188; P:inactivation of MAPK; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR00340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR00387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.


```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 16 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 7) (MAP kinase
DE phosphatase 7) (MKP-7).
GN DUSP16 OR MKP7 OR KIAA1700.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21486429; PubMed=11489891;
RA Masuda K., Shima H., Watanabe M., Kikuchi K.;
RT "MKP-7, a novel mitogen-activated protein kinase phosphatase,
RT functions as a shuttle protein."
RT J. Biol. Chem. 276:39002-39011(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RA MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RT DNA Res. 7:347-355(2000).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB052156; BAB40814.1; -
DR EMBL; AB051487; BAB21791.1; ALT_INIT.
DR HSSP; Q16828; 1MKP.
DR Genew; HGNC:17909; DUSP16.
DR MIM; 607175; -
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0000188; P:inactivation of MAPK; TAS.
DR GO; GO:0045209; P:leptomycin B-sensitive MAPK phosphatase nuc. . .; TAS.
DR GO; GO:0045204; P:MAPK nucleus export; TAS.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR Hydrolase; Nuclear protein.
KW DOMAIN 22 137 RHODANESE.
FT DOMAIN 227 289 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 244 244 BY SIMILARITY.
SQ SEQUENCE 665 AA; 73101 MW; 1BD853FF08460DFP CRC64;
Query Match 7.4%; Score 253.5; DB 1; Length 665;
Best Local Similarity 29.6%; Pred. No. 3.2e-08;
Matches 84; Conservative 44; Mismatches 99; Indels 57; Gaps 9;

```

```

QY 330 SRIFPHLYIGSEWNAANLEELQNRKVTIILMAREI--DNFYPERFTYHNRILMDEBSAQ 387
Db 160 TRILPNLYIGCQRDLNKLKELMQONGIGYVLNASNTPCKPDFIPESH-FLRYVFNDSFCEK 218
QY 388 LLPHWKETHRFIEAPAQGTIVLVHCKMGVSRSAATVLAAYAMKQYECLEQALRHVQELR 447
Db 219 ILPWLKSDYDFIEKAKASNGCVLVHCLAGISRSATIAIAYIMKRMDSLDEAYRFVKEKR 278
QY 448 PIARPNGGLRQLQIYQGITLTSRQSHWEQKVGVSPEEHPAPE----VSTPFPPLP 502
Db 279 PTISPNFNFLGLLDY-----EKKIKNOTGASGPKSKLKLHLKEPNEFPVPA 325
QY 503 EPEGGE-EKVGMEESSQAPKEPFRPRIN-----LRGVM 538
Db 326 VSEGQKSETPLSPPCADSATSEAGQRP-VHPASVPSPVSQPSLLEDSPLVQALSLH 384
QY 539 RSISLEP-----SLESTSETSDMP--EYFSSHESSHE 571
Db 385 LSADRLSDSNKLRKSFSLDIKSVSYSASMAASLHGFSSEDAL 428

RESULT 4
DUS8_MOUSE
ID DUS8_MOUSE STANDARD; PRT; 663 AA.
AC 009112;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 8 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Neuronal tyrosine threonine phosphatase 1).
GN DUSP8 OR NTPPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96311565; PubMed=8733137;
RA Theodosiou A.M., Rodrigues N.R., Nesbitt M.A., Ambrose H.J.,
RA Paterson H., Mclellan-Arnold E., Boyd Y., Leversha M.A., Owen N.,
RA Blake D.J., Ashworth A., Davies K.E.;
RT "A member of the MAP kinase phosphatase gene family in mouse
RT containing a complex trinucleotide repeat in the coding region.";
RL Hum. Mol. Genet. 5:575-684(1996).
CC -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
CC PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
CC SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN AND LUNG.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95518; CA64772.1; -.
DR HSSP; Q16828; IMKP.
DR MGD; MGI:106626; Dusp8.
DR InterPro; IPR000340; DS_Phasphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.

```



```

QY      232 YQERLNSEQSCINEMWTAMADLESIRPPSAEP---GGSSEQEQMOAI---RAELMKVLVD 284
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      236 YDENTNEPSRVMPQOPHIVLESJKREGKEPLVLKGLSSFKQMHENLCDNSLQJQECRE 295
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY      285 VSDLESVTSKEIROALELRIGLPLQCYRDFIDNQMLLLVAQRDRASRIFPHLYIGSEWNA 344
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      296 VGGGASAASSILPQ-----PIPTTPD-IENAEI-----TPILPFLGNEQDA 337
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY      345 ANLEELQRRNVTHILNMAREIDNFYPER--FTYHNVRWLWDEESAOLLPHMKETHRFIEA 402
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      338 QDLDTMQRLNIGYIVNVTTHPLPHYHEKGEJENYKRLPATDSNKQULRQYFEEAFEFIEEA 397
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY      403 RAQGTIVLVHCKMGVSRSAATVLAYAMQYECSLQALRHVQELPFIARPNPGFLRQLOI 462
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      398 HQCGKGLLIHCQAGVSRSAATIVIAYLMKHTMTMTDAYKFKVKRPRITISPLNFMGOLLE 457
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY      463 YQGITASRQSHVWEQKVGCV 483
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db      458 FEEDLNNGVTPRILTPKLMGV 478
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

```

```

RESULT 6
DUS5_RAT
ID_DUS5_RAT      STANDARD;      PRT;      384 AA.
AC      054838;
DT      15-JUL-1999 (Rel. 38, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Dual specificity protein phosphatase 5 (EC 3.1.3.48) (EC 3.1.3.16)
      (MAP-kinase phosphatase CPG21).
GN      DUSP5.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Wistar;
RX      MEDLINE=98364306; PubMed=9699150;
RA      Hevroni D., Ratner A., Bundman M., Lederfein D., Gabarah A.,
RA      Mangelus M., Silverman M.A., Kedari H., Naor C., Kornuc M., Hanoch T.,
RA      Seger R., Theil L.E., Nedivi E., Richter-Levin G., Citri Y.;
RT      "Hippocampal plasticity involves extensive gene induction and multiple
RT      cellular mechanisms.";
RL      J. Mol. Neurosci. 10:75-98(1998).
CC      -1- FUNCTION: THIS PROTEIN SHOWS BOTH ACTIVITY TOWARD TYROSINE-PROTEIN
CC      PHOSPHATE AS WELL AS WITH SERINE/THREONINE-PROTEIN PHOSPHATE (BY
CC      SIMILARITY).
CC      -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC      tyrosine + phosphate.
CC      -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC      phosphate.
CC      -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC      -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC      TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC      -1- SIMILARITY: Contains 1 rhodanese domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/or
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF013144; AAB94858.1; -
DR      HSSP; Q16828; IMKP.
DR      InterPro; IPR000340; DS_phosphatase.
DR      InterPro; IPR01763; Rhodanese-like.
DR      InterPro; IPR000387; TYR_phosphatase.
DR      Pfam; PF00782; DSPC; 1.
DR      Pfam; PF00581; Rhodanese; 1.
DR      SMART; SM00195; DSPC; 1.

```

DR	SMART; SM00450; RHOD; 1.	
DR	PROSITE; PS50206; RHODANESE_3; 1.	
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.	
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.	
DR	PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.	
KW	Hydrolase; Nuclear protein.	
FT	DOMAIN 19 141	RHODANESE.
FT	DOMAIN 180 384	PROTEIN-TYROSINE PHOSPHATASE.
FT	DOMAIN 53 74	NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	DOMAIN 79 82	POLY-GLY.
FT	ACT SITE 263 263	BY SIMILARITY.
SQ	SEQUENCE 384 AA; 42094 MW; 5644069B8D348700 CRC64;	

Query Match	6.8%;	Score 233;	DB 1;	Length 384;
Best Local Similarity	31.2%;	Pred. No. 2.7e-07;		
Matches 67;	Conservative 30;	Mismatches 84;	Indels 34;	Gaps 5;

[illegible]

```

RESULT 7
DUSA_MOUSE
ID DUSA_MOUSE STANDARD; PRT; 483 AA.
AC Q9ES50; Q9CZY9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 10 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 5) (MAP kinase
DE phosphatase 5) (MKP-5).
GN DUSP10 OR MKP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20515587; PubMed=11060451;
RA Masuda K., Shima H., Kikuchi K., Watanabe Y., Matsuda Y.;
RT "Expression and comparative chromosomal mapping of MKP-5 genes
RT DUSP10/Dusp10.";
RL Cytogenet. Cell Genet. 90:71-74(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

```

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyko-Oka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Involved in the inactivation of MAP kinases. Has a
CC specificity for the MAPK11/MAPK12/MAPK13/MAPK14 subfamily (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AB037908; BAB17680.1; -;
DR EMBL, AK011995; BAB27966.1; -;
DR HSSP, Q16828; 1MKP.
DR MGD, MGI:1927070; Dusp10.
DR InterPro, IPR000340; DS_phosphatase.
DR InterPro, IPR001763; Rhodanese-like.
DR InterPro, IPR000387; TYR_phosphatase.
DR Pfam, PF00782; DSpC; 1.
DR Pfam, PF00581; Rhodanese; 1.
DR SMART, SM00195; DSpC; 1.
DR SMART, SM00450; RHOD; 1.
DR PROSITE, PS50206; RHODANESE_3; 1.
DR PROSITE, PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE, PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE, PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase; Nuclear protein.
FT DOMAIN 169 286 RHODANESE.
FT ACT SITE 386 456 PROTEIN-TYROSINE PHOSPHATASE.
FT CONFLICT 409 409 BY SIMILARITY.
FT CONFLICT 48 48 A -> T (IN REF. 2).
FT CONFLICT 113 113 P -> S (IN REF. 2).
FT CONFLICT 386 386 R -> Q (IN REF. 2).
SQ SEQUENCE 483 AA; 52530 MW; 3551562355B46F89 CRC64;
Query Match 6.8%; Score 232.5; DB 1; Length 483;
Best Local Similarity 26.2%; Pred. No. 3.9e-07;
Matches 77; Conservative 45; Mismatches 113; Indels 59; Gaps 8;

QY 232 YQERLNSEQSCINWTAMADLESIRPPSAEP---GGSSHQEQMEQAI---RAELMKVLD 284
DB 237 YDENTNEPRVTPSPQPLHIVLESIRKEGKEPLVLKGLSSFKONHGNLCONSLOQECRE 296
QY 285 VSDLESVTSKEIRQALELRLGLPLQOYRDFIDNOMLLVAQRDRASRIFFHYLGSEWNA 344
DB 297 VGGASASASSMLPQSVPT-----TPDIENAEI-----TPIPLFLFNGEODA 338
QY 345 ANLEELQNRVTHILMAREIDNYPBER--FTYHNVRIMDEESAQLLPHWKETHRFIEBA 402
DB 339 QDLDTMQRNLNIGVINVTTHLPLHYHEKGLFNKRLPATDSNKNLRRYFEBAFEFIEBA 398
QY 403 RAQGTHTLVHCKMGVSRSAATVLAAMKQYECLEQALRHVQELRPIARPNGFLRQLOI 462
DB 399 HQCGKGLLHCQAGVSRSAIVIAVMKHTMTDAYKFKVKGKRPITISPLNFMGQILLE 458
QY 463 YQGITLTSRQSHVWEQKVGVSPFEHPAPEVSTPFPPLPEPEGGEKQVGM 516

Db 459 FE-----EDLNNGVT-----PRILTP-----KLMGME 480
RESULT 8
DUSF_HUMAN STANDARD; PRT; 295 AA.
AC Q9H1R2; Q8N826; Q9BX24;
DT 28-FEB-2003 (Rel. 41, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 15 (EC 3.1.3.48) (EC 3.1.3.16).
GN DUSP15 OR C20ORF57.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Testis;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.,
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11760052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leveaesiainho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLeay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
RA Skuon C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Wittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9H1R2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9H1R2-2; Sequence=VSP_007292, VSP_007293;
CC Note=Derived from EST data;
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- CAUTION: Ref.2 sequences differ from that shown due to erroneous
CC gene model prediction.
CC -1- CAUTION: Although assigned as two separate genes (C20orf57 and
CC DUSP15), it is probable that C20orf57 does not exist by itself and
CC is a part of the DUSP15 gene.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK097430; BAC05048.1; -;
DR EMBL; AL160175; CAC10008.2; ALT_SEQ.
DR EMBL; AL160175; CAC28981.1; ALT_SEQ.
DR EMBL; BM554314; -; NOT_ANNOTATED_CDS.
DR HSSP; P51452; LVHR.
DR Genew; HGNC:16236; DUSP15.
DR Genew; HGNC:16190; C20orf57.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SMO0195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolyase; Alternative splicing.
FT DOMAIN 62 132 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 85 85 BY SIMILARITY.
FT VARSPLIC 143 232 GARRTSKTSQAQCPMTSATCLLAARVALLSAALVREATG
FT RTAQCRLSPRAAERLGPPIVAAGWSPDPKYQICLFCG
FT EEDPGPTQ -> LRRLERFGESPFRDEELRALPLCKR
FT CRGGSATSSAGPHSAASEGTIVQLVPTPREAHPPLPL
FT ARVKQTFSCLPCLSRKSGK (in isoform 2).
FT /FTId=VSP_007292.
FT VARSPLIC 233 295 Missing (in isoform 2).
FT /FTId=VSP_007293.
SQ SEQUENCE 295 AA; 31881 MW; 28F8A687ECB5C219 CRC64;
Query Match 6.7%; Score 229; DB 1; Length 295;
Best Local Similarity 27.2%; Pred. No. 3.4e-07;
Matches 71; Conservative 39; Mismatches 119; Indels 32; Gaps 6;

DB 64 FKECINFIHCRLNGNCNCFAGISRTTITVAVMTVTGLGWRDYLAIKATRIPIAN 123
QY 452 PNPGRQLQIYQGITLTAQRSHWWEQKVGVSPEEHP-----APEVETPPPLPPEPEG 506
DB 124 PNPGRQLQLEFVGWASSQKARHRTSKTSGAQCPEMTSATCLLAARVALLSAALVREATG 183
QY 507 -----GGEKVVGWEESSQA-----PK-----EEDGPRPRINLRGVMRSI 541
DB 184 RTAQCRLSPRAAERLGPPIVAAGWSPDPKYQICLFCGEEDPGFTQHPKEQLIMADV 243
QY 542 SL-LEPSLELESTSETSDMPE 561
DB 244 QVQLRPGSSSCTLSASTERD 264
RESULT 9
DUS7_MOUSE STANDARD; PRT; 320 AA.
ID DUS7_MOUSE
AC Q91Z46;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Dual specificity protein phosphatase 7 (EC 3.1.3.48) (EC 3.1.3.16).
GN DUSP7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; Pubmed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC010207; AAH10207.1; -;
DR MGD; MGI:2387100; Dusp7.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.

DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase.
FT ACT_SITE 209 277 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 320 AA; 35258 MW; 2AAD2723E27DAE53 CRC64;
Query Match 6.6%; Score 226.5; DB 1; Length 320;
Best Local Similarity 25.0%; Pred. No. 5.3e-07;
Matches 91; Conservative 43; Mismatches 117; Indels 113; Gaps 12;
QY 152 PDSSPSCCTGLVPLWSDTQVYLDG-----DGGFVSSTSGQSRIFFXISIQTWATLQ 205
DB 46 PEPGAPASVGLLLQKLRD-----DGCQAYVLGGFN-----KQTEYS--- 84
QY 206 VLHQACEALGSGLVPGSGALTWASHYQERLNSQSCINWTAMADLESRLPPSAEPGGS 265
DB 85 ---EHCEITVDSSSSPSGSP-----TSVLGLGGLRISDCSDGE 121
QY 266 SEQOMEQAIRAELWKLVDVSDLESVTSKEIRQALERLGLPLQYRDFIDNQLLVAQ 325
DB 122 SDRLEPSATE-----SDGSPVPSQ-----PAFPVQ----- 148
QY 326 RDRASRIFFHLVYLGSEWNAANLELQNRVTHILNAREIDNYPE--RTYHNRVLMDE 383
DB 149 -----IDPYLYLGCAKSTNLDVLGKYGIKILNVTPLNPAFHEGGEFTYKQIPISDH 202
QY 384 ESAQLPHEKETHRFIEARAQTHVLVHCKMGVRSATVLAAMQYECLEQALRHV 443
DB 203 WSQNLQCFPEPAISFIDEARSKCGVLVHCLAGISRSVTVAAYLMQKNLSLNDAYDFV 262
QY 444 QELRPIAPBNPGFLRQLQYQGITASRQSHWEQKVGVSPEHPAPE---VSTP--- 496
DB 263 KRKKSNIENFNFMGLD-----ERTLGLSSPCDNHAPSEQLYFSTPTNH 309
QY 497 --FP 498
DB 310 NLF 313
RESULT 10
DUS9_HUMAN STANDARD; PRT; 384 AA.
ID DUS9_HUMAN
AC Q99956;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 9 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase 4) (MAP kinase phosphatase 4) (MKP-4).
GN DUSP9 OR MKP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97184169; PubMed=9030581;
RA Muda M., Boschart U., Smith A., Antonsom B., Gillieron C.,
RA Chabert C., Camps M., Martinou I., Ashworth A., Arkinstall S.,
RT "Molecular cloning and functional characterization of a novel
RT mitogen-activated protein kinase phosphatase, MKP-4.";
RL J. Biol. Chem. 272:5141-5151(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Platzner M.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INACTIVATES MAP KINASES. HAS A SPECIFICITY FOR THE ERK
CC FAMILY.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein

CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y08302; CAA69610.1; -.
CC EMBL; U52111; -; NOT_ANNOTATED_CDS.
CC HSSP; Q16828; MKP.
CC DR GeneW; HGNC:3076; DUSP9.
CC DR MIM; 300134; -.
CC DR GO; GO:0005737; Cytoplasm; TAS.
CC DR GO; GO:0005634; Cnucleus; TAS.
CC DR GO; GO:0000188; P:inactivation of MAPK; TAS.
CC DR GO; GO:0007254; P:JNK cascade; TAS.
CC DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
CC DR InterPro; IPR000340; DS_phosphatase.
CC DR InterPro; IPR001763; Rhodanese-like.
CC DR Pfam; PF00782; DSPC; 1.
CC DR SMART; SM00195; DSPC; 1.
CC DR SMART; SM00450; RHOD; 1.
CC DR PROSITE; PS50206; RHODANES_3; 1.
CC DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
CC DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
CC DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase.
FT DOMAIN 18 139 RHODANES.
FT ACT_SITE 203 384 PROTEIN-TYROSINE PHOSPHATASE.
FT ACT_SITE 290 290 BY SIMILARITY.
SQ SEQUENCE 384 AA; 41867 MW; F8598CA95AB379B7 CRC64;
Query Match 6.6%; Score 225.5; DB 1; Length 384;
Best Local Similarity 30.0%; Pred. No. 7.8e-07;
Matches 68; Conservative 30; Mismatches 92; Indels 37; Gaps 5;
QY 284 DVSDLESVTSKEIRQA-----LERLGLPLQYRDFIDNQLLVAQRDRASRI 332
DB 170 DCSDAESADRDMSGCLDSEGATPPVGLRASFPVQ-----I 207
QY 333 FPHLYLGSEWNAANLELQNRVTHILNAREIDNYPE--FTYHNRVLMDESAQLLP 390
DB 208 LPNLYLGASDASANLESLAKGIRYILNVTPLNPFENKGDHYKQIPISDHWSQNLRS 267
QY 391 HMKETHRFIEARAQTHVLVHCKMGVRSATVLAAMQYECLEQALRHVQELRPIA 450
DB 268 FFPEAIEFIDALSQNCGLVHCLAGVSRSVTVVAAYLMQKHLSLNDAYDLVKRKSNI 327
QY 451 RPNPGFLRQLQYQGITASRQSHWEQKVG--VSPEHPAPEVSTP 496
DB 328 SPNFMGLDIDFERSRL--EERHSGEQSGGQASASNPSPFTTP 373
RESULT 11
DUS5_HUMAN STANDARD; PRT; 384 AA.
ID DUS5_HUMAN
AC Q16690; Q12997;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 5 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase hvh3).
GN DUSP5 OR VH3.


```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95138103; PubMed=7836374;
RA Kwak S.P., Dixon J.E.;
RT "Multiple dual specificity protein tyrosine phosphatases are
RT expressed and regulated differentially in liver cell lines.";
RL J. Biol. Chem. 270:1156-1160(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=95050849; PubMed=7961985;
RA Ishibashi T., Bottaro D.P., Michieli P., Kelley C.A.,
RA Aaronson S.A.;
RT "A novel dual specificity phosphatase induced by serum stimulation
RT and heat shock.";
RL J. Biol. Chem. 269:29897-29902(1994).
RN [3]
RP REVISIONS.
RA Bottaro D.P.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DISPLAYS PHOSPHATASE ACTIVITY TOWARD SEVERAL SUBSTRATES.
CC THE HIGHEST RELATIVE ACTIVITY IS TOWARD ERK1.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- SIMILARITY: Contains 1 rhodanese domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U16996; AAB06261.1; -.
DR EMBL; U15932; AAA64693.2; -.
DR PIR; I38890; I38890.
DR HSSP; Q16828; LMKP.
DR Genew; HGNC:3071; DUSP5.
DR MIM; 603069; -.
DR GO; GO:0004725; F:protein tyrosine phosphatase activity; TAS.
DR GO; GO:0006470; P:protein amino acid dephosphorylation; TAS.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR001763; Rhodanese-like.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR Pfam; PF00581; Rhodanese; 1.
DR SMART; SM00195; DSPC; 1.
DR SMART; SM00450; RHOD; 1.
DR PROSITE; PS50206; RHODANESE_3; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydroxylase; Nuclear protein.
FT DOMAIN 19 141 RHODANESE.
FT DOMAIN 180 384 PROTEIN-TYROSINE PHOSPHATASE.
FT DOMAIN 53 74 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 79 82 POLY-Gly.
FT ACT_SITE 263 263 BY SIMILARITY.
FT CONFLICT 9 11 RQL -> GHV (IN REF. 2).
FT CONFLICT 71 71 A -> R (IN REF. 2).
FT CONFLICT 105 106 AR -> F (IN REF. 2).
SQ SEQUENCE 384 AA; 42107 MW; D2B726F7C0414306 CRC64;

```

	Query Match	6.6%;	Score 224;	DB 1;	Length 384;
	Best Local Similarity	27.7%;	Pred. No. 9.6e-07;		
	Matches 62;	Conservative 40;	Mismatches 84;	Indels 38;	Gaps 5
QY	287 DLESVTSKEI--RQALELRIGLPL--QQYRDFIDNQMLLVQAQRDRASRIFFPHLYLGSEW	342	:	:	:
Dd	142 DVKPISEKIESERALISQCCKPWNVSYPAYD-----QGPEVLIPFLYLGSAY	192	:	:	:
QY	343 NAANLLEPQNRNVTTHILNMAREIDNFYPERFTYHNVRWLWDEESAQLLPWKETHRFTEAA	402	:	:	:
Dd	193 HASKCEFLANLHITALLNVSRTSEACMTHLHYKWIIVEDSHITADISSHFQEAIDFIDCV	252	:	:	:
QY	403 RAQSTHYLVHCKMGVSRSAATVLAAMKQYECSLAQALRHVOELRPRIARPNGFLRQLCI	462	:	:	:
Dd	253 REKGKGVLVHCAGISRSPITCMAYLMKTQFRLEKAFDYIKQRSMVSFNFGFMGQLLQ	312	:	:	:
QY	463 YQGILTASRQSHWEQKVGVSPBEHPAPEVSTPPFPPLPPEPEG	506	:	:	:
Dd	313 YESELTP-----STP-NPQPSCCG	331	:	:	:

```

RESULT 12
DUST7_RAT
ID DUST7_RAT STANDARD; PRT; 280 AA.
AC 063340;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 7 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Dual specificity protein phosphatase MKP-X) (Fragment).
GN DUSP7 OR MKPX.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Neuron;
RX MEDLINE=96224012; PubMed=8626780;
RA Muda M., Boschart U., Dickinson R., Martinou J.C., Martinou I.,
RA Camps M., Schlegel W., Arkinstall S.;
RT "MKP-3, a novel cytosolic protein-tyrosine phosphatase that
RT exemplifies a new class of mitogen-activated protein kinase
RT phosphatase.";
RL J. Biol. Chem. 271:4319-4326(1996).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC -1- tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X94186; CAA63896.1; -.
DR HSSP; Q16828; MKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydrolyase.
PT NON_TER 1 1

```

[illegible]

```

DR   Genew; HGNC:3073; DUSP7.
DR   MIM; 602749; -.
DR   GO; GO:0005737; C:cytoplasm; ISS.
DR   GO; GO:0004725; F:protein tyrosine phosphatase activity; ISS.
DR   GO; GO:0000188; P:inactivation of MAPK; ISS.
DR   InterPro; IPR000340; DS_phosphatase.
DR   InterPro; IPR000387; TYR_phosphatase.
DR   Pfam; PF00782; DSPC; 1.
DR   SMART; SM00195; DSPC; 1.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; FALSE_NEG.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
DR   HydroLase.
KM   DOMAIN 209 277 PROTEIN-TYROSINE PHOSPHATASE.
FT   ACT_SITE 232 232 BY SIMILARITY.
SQ   SEQUENCE 320 AA; 35278 MW; D387F6BEFBA9213C CRC64;

Query Match 6.4%; Score 217.5; DB 1; Length 320;
Best local similarity 24.2%; Pred. No. 1.9e-06;
Matches 84; Conservative 45; Mismatches 125; Indels 93; Gaps 10

QY 152 PDSSSPECTLGLVLPWMSDTQVYLDG-----DGFSVTSGGQSRIFKPISIQTMMATLQ 205
DB 46 PEPGAPASVGLILLQKLRD-----DGCQAYYLGGFN-----KQTEYS--- 84

QY 206 VLHQACEALIGSGLVPGGSALTWASHYQERLNSQSCLENWTAMADLESILRPPSAEPGGS 265
DB 85 --EHCEINVDSSSSPSSSP-----TSVLGLGLRISSDCSDGE 121

QY 266 SEQEQMEQAIRAELWKVLVDSDLESVTSKEIRQALELRGLPLQQRDFIDNQMLLLVAQ 325
DB 122 SDRELPSATE-----SDGSPVPSSQ-----PAFPVQ----- 148

QY 326 RDRASRIPEHLVYLGSEWNAANIELQRNVTILNMAREIDNFYPE--RFTYHNVKLWDE 383
DB 149 -----ILPYLYLGCAKDSINLDVLGKYGIKYLNTVNLNPAFEGEGEFTYKQIPISDH 202

QY 384 ESAQLPFWKETHRFIEARAQGTIVLVCKMGVSRSAATVLAAMKQYECSLEQALRHV 443
DB 203 WSONLSQEPPEAISIFIDEARSKKCGVLVHCLAGISRSVTVTVAYLMQKNLSINDAYDFV 262

QY 444 QELRPIARPENGFILROLQIYQGILITASR--QSHWWEQYKGVSPPEEH 488
DB 263 KRKKSNIENFNFMGQLDFERTLGLSSPCDNHASSQQLYFSTPTNH 309

RESULT 14
DUS4_RAT
ID DUS4_RAT STANDARD; PRT; 395 AA.
AC Q62767;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dual specificity protein phosphatase 4 (EC 3.1.3.48) (EC 3.1.3.16)
DE (Mitogen-activated protein kinase phosphatase-2) (MAP kinase
DE phosphatase-2) (MKP-2).
DE GN DUSP4 OR MKP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TITISUE=Pheochromocytoma;
RX MEDLINE=95301550; PubMed=7782322;
RA Mista-Press A., Rim C.S., Yao H., Roberson M.S., Stork P.J.S.;
RA "A novel mitogen-activated protein kinase phosphatase. Structure,
RT expression, and regulation.";
RL J. Biol. Chem. 270:14587-14596(1995).
CC -!- FUNCTION: REGULATES MITOGENIC SIGNAL TRANSDUCTION BY
CC DEPHOSPHORYLATING BOTH THR AND TYR RESIDUES ON MAP KINASES
CC ERK1 AND ERK2 (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein

```

CC tyrosine + phosphate.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED AT MODERATE LEVELS IN NEARLY ALL
CC TISSUES AND CELLS INCLUDING BRAIN, SPLEEN, AND TESTES WITH THE
CC HIGHER EXPRESSION IN THE HEART AND LUNG AND LOWER EXPRESSION IN
CC SKELETAL MUSCLE AND KIDNEY. UNDETECTABLE IN LIVER. EXPRESSED IN
CC MANY AREAS OF THE BRAIN WITH VERY STRONG EXPRESSION IN THE
CC HIPPOCAMPUS, PIRIFORM CORTEX, AND THE SUPRACHIASMATIC NUCLEUS.
CC -1- INDUCTION: BY MITOGENS AND BY STRESS.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -1- SIMILARITY: Contains 1 rhodanese domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC	-----	
DR	EMBL; U23438; AAC52493.1; -.	
DR	HSSP; Q16828; IMKP.	
DR	InterPro; IPR000340; DS_phosphatase.	
DR	InterPro; IPR001763; Rhodanese-like.	
DR	InterPro; IPR000387; TYR_phosphatase.	
DR	Pfam; PF00782; DSPC; 1.	
DR	Pfam; PF00581; Rhodanese; 1.	
DR	SMART; SM00195; DSPC; 1.	
DR	SMART; SM00450; RHOD; 1.	
DR	PROSITE; PS50206; RHODANESE_3; 1.	
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.	
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.	
DR	PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.	
KW	Hydrolase; Nuclear protein.	
FT	DOMAIN 42 160	RHODANESE.
FT	DOMAIN 198 395	PROTEIN-TYROSINE PHOSPHATASE.
FT	ACT SITE 281 281	BY SIMILARITY.
SC	SEQUENCE 395 AA; 43187 MW; A90EFFFD378A050FD CRC64;	

Query Match	6.3%;	Score 215.5;	DB 1;	Length 395;
Best Local Similarity	22.7%;	Pred. No. 3.3e-06;		
Matches	87;	Conservative	54;	Mismatches 130;
				Indels 113;
				Gaps 10;

QY	176	DGDGGSFVTSGGGSRIFKPIISIQIMWATLOVLHQA	CEALGSLVPGGSALTW-----	228
		: : : :	: : : :	
Db	22	DENGGTAGSSGGG-----	-----HGAL-----	GILSGKCLLLDCRPFLLA 56
QY	229	--ASHYQERLNSEQSCINEWTA--MADLESIRPPSAEPGSS	EQEOMEQAIRAELMKVLD 284	
		: : : :	: : : :	
Db	57	HSAGYIRGSVAVRCNTIVRRRAKGSVSLQILP-----	AAEEVRLRRSLGYSAVI 107	
QY	285	VSDLESVTSKEIRQALELRLGLPLQYRDFID-NQMLLVAQDRAS-----	330	
		: : : :	: : : :	
Db	108	VYDERSPRAESLRE--DSTVSLVQALRMAERTD	ICLLKGYERFSSEYEPFCSKTRAL 165	
QY	331	-----	-----RIFPHLYLQSEMNANLEELQNRVTH 357	
Db	166	AAIRPPVPPESTINESL DLGSSCGTPLHDQGPVEILP	FLYLGSAYYHAARDMDALGTTA 225	
QY	358	ILNMAREIDNFPYPERFTYHNVRLEWDEESAQLLPHWKETH	RPIEARAQGTHTVLVHCKMGV 417	
		: : : :	: : : :	
Db	226	LLNVSSDCPNHFECHGYQYKCIPVEDNHKADISSW	MEALEYIDAVKDCRGRLVHQAIGI 285	
QY	418	SRSAAATVLAYAMKQYECSLQALRHQELRPIARPNGFLRO-LQIYQ	GILTASRQSHW 476	
		: : : :	: : : :	
Db	286	SRSATILCLAYLMKKRVRLEEAFFVKQRSLISPNFS	FMGQLLFQESQVLTTSS----- 339	
QY	477	EOKVGVSPDEHPAPEVSTPFPPL 500		
		: : : :		
Db	340	-----	-----CAAEFAASPGPL 351	

```

RESULT 15
VHPI_CAEEL
ID VHPI_CAEEL STANDARD; PRT; 619 AA.
AC Q10038;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein-tyrosine phosphatase vhp-1 (EC 3.1.3.48).
GN VHP-1 OR F08B1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoea; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Chisoe S.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY. DUAL SPECIFICITY SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; U23178; AAC46719.1; -.
DR PIR; T15969; T15969.
DR HSSP; Q16828; IMKP.
DR WormPep; F08B1.1; CE01899.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KM Hydrolase.
FT DOMAIN 200 269 PROTEIN-TYROSINE PHOSPHATASE.
FT FT 92 95 POLY-SER.
FT DOMAIN 351 354 POLY-SER.
FT FT 465 472 POLY-SER.
FT DOMAIN 483 488 POLY-SER.
FT ACT SITE 224 224 BY SIMILARITY.
SQ SEQUENCE 619 AA; 66354 MW; 369E326F615D0529 CRC64;

```

Query Match	6.2%;	Score 210.5;	DB 1;	Length 619;
Best Local Similarity	20.9%;	Pred. No. 1.2e-05;		
Matches 113;	Conservative 66;	Mismatches 198;	Indels 163;	Gaps 18;

QY	268	QEONEQAIRAEALWKYLD---	VSDLESVTSKEIRQALEIRL-----	GLPL-----	308
		: : : :	:	:	:
		: : : :	:	:	:
Db	40	RERLEDTDKFRSVMVLEEODPVSSLPTFP	PAKKFGLKQLTLTSSPTNSSPPISSSSPTNG	99	
QY	309	-----QQYRDFID-----	NOMLLVAQRBRASRIFFPHLYLGSEMANA	EEL	350
		:	:	:	:
		:	:	:	:
Db	100	GFKQFAQQYPOLCESSEGMRRLPOSLSQPCLSQPTGDI	TILITPNITYLGSQIDSDETML	159	
QY	351	QRNRVTIILNMA-----	REIDNFYPERFTYHNVRVLMDDESAOQLPHWKE	THFIEA	401
		: : : :	:	:	:
		: : : :	:	:	:
Db	160	DAIDISVINLSMTCPSVCIKEDKNFM-----	RIPVNDSYQEKLSPYFMAVEFLEK	212	
QY	402	AAAGTHTLVHCKMGVSRSATVLAAYAMKQVECSLEQALRHVOELRP	IARPNPGFLROLQ	461	
		: : : :	:	:	:
		: : : :	:	:	:
Db	213	CRAKGKCLIHCLAIISSPTIAISYIMRYMKMGSDDAYRVVKERRPSI	SPENNEWGOLL	272	

Tue Jan 20 06:36:39 2004

us-09-955-732-2.rsp

Page 12

```

QY      462 IYOGIL-----TASR-----QSHVWEQKV-----GGVSPHEEPA-- 490
      | : |
      | : |
Db      273 EYENVLIKDHVLDYNQASRPHRMIDYGGPSDLCPPKVPKASNCVFPFGSTHDESSPSSP 332
      | : |
      | : |
QY      491 -----PEVST-----PEPP-----LPEE 503
      || |
      || |
Db      333 SVSEGSAASEFEPTSSSAASSSSTASAPPSPMPTSEOGTSGTVNVNNGKRMNTMDGLPHR 392
      | : |
      | : |
QY      504 PEGGEEKVVGMEESO-AAPKEE-----PGPR-----PRINRGVM 538
      | : |
      | : |
Db      393 PKALGLPFRIGTSVAELPSPSTELSRUSFNGPEAIAPSTPLINTPCFNPIIHYAASS 452
      | : |
      | : |
QY      539 RSISLLEPSTLESTSETSDMPEV-FSSHSSHEEPL--QEPFQLARTKGQGYDRGPQ 594
      | : |
      | : |
Db      453 REVILTLTPAASSSSSSSTSESPSFDPSSFESSSSSSSIVENEFFASTVPAGSSSISTPS 512
      | : |
      | : |
QY      595 PALKSRQSVVTLQGSAAVVANRTQAFQEQEQGGQGGQGGEPCTISSTPRFRKVVRQASVHDSG 654
      | : |
      | : |
Db      513 ---GSGSTPASASSSSAASRCRMKGFFKVFVKKAPASTSTPASSTPGTSRAARECRRSSG 569

```

Search completed: January 15, 2004, 06:49:52
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: January 15, 2004, 06:41:46 : Search time 40 Seconds
(without alignments)
4251.417 Million cell updates/sec

Title: US-09-955-732-2
Perfect score: 3412
Sequence: 1 MALVTVSRSPPGSGASTPVG.....RFRKVVROASVHDSGEEGEA 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_23:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_virus:*
 - 16: sp_bacteriaph:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3404	99.8	659	4 Q8TE77	Q8TE77 homo sapien
2	2627	77.0	649	11 Q8K330	Q8K330 mus musculu
3	2623	76.9	513	4 Q8N9L8	Q8N9L8 homo sapien
4	2422	71.0	471	4 Q8WYL0	Q8WYL0 homo sapien
5	1884.5	55.2	394	4 Q9NMZ7	Q9NMZ7 homo sapien
6	1564	45.8	299	4 Q9BQ20	Q9BQ20 homo sapien
7	1172	34.3	692	4 Q8WYL4	Q8WYL4 homo sapien
8	1172	34.3	1049	4 Q8WYL5	Q8WYL5 homo sapien
9	1136	33.3	703	4 Q8N9A7	Q8N9A7 homo sapien
10	1133	33.2	464	11 Q8C241	Q8C241 mus musculu
11	1089.5	31.9	449	4 Q8WYL2	Q8WYL2 homo sapien
12	940.5	27.6	1045	5 Q9NMY1	Q9NMY1 drosophila
13	918.5	26.9	1193	5 Q8IMU8	Q8IMU8 drosophila
14	581.5	17.0	141	4 Q9NV45	Q9NV45 homo sapien
15	416.5	12.2	738	4 Q9P2P8	Q9P2P8 homo sapien
16	405	11.9	127	4 Q8TDB5	Q8TDB5 homo sapien

17	366.5	10.7	195	4 Q96H36	Q96H36 homo sapien
18	358.5	10.5	195	4 Q8WYL1	Q8WYL1 homo sapien
19	322	9.4	202	4 Q96F40	Q96F40 homo sapien
20	310.5	9.1	148	4 Q8WYL3	Q8WYL3 homo sapien
21	288	8.4	198	10 Q9ZR37	Q9ZR37 arabidopsis
22	288	8.4	198	10 Q9LUG6	Q9LUG6 arabidopsis
23	269.5	7.9	220	11 Q9D6P6	Q9D6P6 mus musculu
24	269.5	7.9	220	11 Q8K4T5	Q8K4T5 mus musculu
25	269.5	7.9	220	11 Q99N12	Q99N12 mus musculu
26	268	7.9	660	11 Q920R2	Q920R2 mus musculu
27	268	7.9	677	11 Q99MG6	Q99MG6 mus musculu
28	258	7.6	662	4 Q8IVT8	Q8IVT8 homo sapien
29	253.5	7.4	616	4 Q8N5T1	Q8N5T1 homo sapien
30	250.5	7.3	162	11 Q9CRR3	Q9CRR3 mus musculu
31	247.5	7.3	665	4 Q96N49	Q96N49 homo sapien
32	246.5	7.2	227	5 Q95XK5	Q95XK5 caenorhabdi
33	242.5	7.1	196	10 Q8LPY0	Q8LPY0 oryza sativ
34	239	7.0	1042	4 Q9C0D8	Q9C0D8 homo sapien
35	233.5	6.8	483	11 Q9CZY9	Q9CZY9 mus musculu
36	233.5	6.8	483	11 Q8R3L3	Q8R3L3 mus musculu
37	233	6.8	622	11 Q99MG5	Q99MG5 mus musculu
38	232.5	6.8	167	10 Q9M8K7	Q9M8K7 arabidopsis
39	232.5	6.8	483	11 Q9ESS0	Q9ESS0 mus musculu
40	232.5	6.8	943	10 Q8L4Z7	Q8L4Z7 oryza sativ
41	232	6.8	184	11 Q99N11	Q99N11 mus musculu
42	229	6.7	295	4 Q8N826	Q8N826 homo sapien
43	228.5	6.7	184	4 Q9NWR4	Q9NWR4 homo sapien
44	227	6.7	177	11 Q9CSL5	Q9CSL5 mus musculu
45	227	6.7	1543	5 Q9GV71	Q9GV71 dictyosteli

ALIGNMENTS

RESULT 1									
ID	Q8TE77	PRELIMINARY;	PRT;	659	AA.				
AC	Q8TE77;								
DT	01-JUN-2002	(TREMblrel. 21, Created)							
DT	01-JUN-2002	(TREMblrel. 21, last sequence update)							
DT	01-MAR-2003	(TREMblrel. 23, last annotation update)							
DE	Hypothetical protein FLJ23852.								
OS	Homo sapiens (Human).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX	NCBI_TaxID=9606;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,								
RA	Suzuki Y., Odayashi M., Nishi T., Shibahara T., Tanaka T.,								
RA	Nakamura Y., Iisgai T., Sugano S.;								
RT	"NEDO human cDNA sequencing project.";								
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AK074432; BAB85080.1; -.								
DR	HSSP; Q16828; IMKP.								
DR	InterPro; IPR000340; DS_phosphatase.								
DR	InterPro; IPR000387; TYR_phosphatase.								
DR	Pfam; PF00782; DSPC; 1.								
DR	SMART; SM00195; DSPC; 1.								
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.								
DR	PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.								
KW	Hypothetical protein.								
SQ	SEQUENCE 659 AA; 72935 MW; 0D96E36F1FE81D3B CRC64;								
Query Match									
Best Local Similarity 99.8%; Score 3404; DB 4; Length 659;									
Matches 658; Conservative 0; Mismatches 1; Indels 0; Gaps 0									
QY	1 MALVTVSRSPPGSGASTPVGPDQAVQRRSRLQRRQSAVLRGAVLGLQDGGNDDAEA 60								
DB	1 MALVTVSRSPPGSGASTPVGPDQAVQRRSRLQRRQSAVLRGAVLGLQDGGNDDAEA 60								
QY	61 SSEPTKAPSEELHGDQTFQGGSGSPQKEQGRQLHLMVQLLRPDQIRLAQLAQLAP 120								


```

Db      61 SSEPTEKAPSEELHGDQDTEFGGQSOSPOKOEQRQHLHMVQLRPQDDIRLAAQLEAP 120
QY      121 RPPRLRLYLAVSTREGEGLSQDETVLLGVDFPDSSSPSCTGLVPLPMSDTQVYLDGDG 180
Db      121 RPPRLRLYLAVSTREGEGLSQDETVLLGVDFPDSSSPSCTGLVPLPMSDTQVYLDGDG 180
QY      181 FSVTSGGQSRIFKPISIQTMATLOVLHQAACEALGSLVPGGSALTWASHYQERLNSE 240
Db      181 FSVTSGGQSRIFKPISIQTMATLOVLHQAACEALGSLVPGGSALTWASHYQERLNSE 240
QY      241 SCLNWTAMADLESRLPPSAEPGSSSEQOEQMEQAIIRAEIMKVLVDVSDLESVTSKEIRQA 300
Db      241 SCLNWTAMADLESRLPPSAEPGSSSEQOEQMEQAIIRAEIMKVLVDVSDLESVTSKEIRQA 300
QY      301 ELRLGLPQQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQNRVTHILN 360
Db      301 ELRLGLPQQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQNRVTHILN 360
QY      361 MAREIDNFEYPERFTYHNVRWLWDEBSAQLLPHWKETHRFIEARAQGTHTVLVHCKMGVSR 420
Db      361 MAREIDNFEYPERFTYHNVRWLWDEBSAQLLPHWKETHRFIEARAQGTHTVLVHCKMGVSR 420
QY      421 AATVLAYAMKQYECSEALRHVQELRPIARPNGFLRQLQIYQIGILTASQSHWEQKV 480
Db      421 AATVLAYAMKQYECSEALRHVQELRPIARPNGFLRQLQIYQIGILTASQSHWEQKV 480
QY      481 GGVSPPEHPAPEVSTPFPPLPPEPEGGGEEKVGMESQAAPKEEPGRPRINLRGVMS 540
Db      481 GGVSPPEHPAPEVSTPFPPLPPEPEGGGEEKVGMESQAAPKEEPGRPRINLRGVMS 540
QY      541 ISLLEPSLEESTSETSDMPEVFSHSSHEEPLQFPQLARTKGQGVDRGPQALKS 600
Db      541 ISLLEPSLEESTSETSDMPEVFSHSSHEEPLQFPQLARTKGQGVDRGPQALKS 600
QY      601 QSVVTLOGSAVVANRTQAFQEQEQGQGGQGEPCISSTPRFRKVVRAQASVDSGEEGEA 659
Db      601 QSVVTLOGSAVVANRTQAFQEQEQGQGGQGEPCISSTPRFRKVVRAQASVDSGEEGEA 659

```

RESULT 2

```

Q8K330 PRELIMINARY; PRT; 649 AA.
ID Q8K330;
AC Q8K330;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to slingshot 3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028922; AAH28922.1;
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 649 AA; 72227 MM; DF8628B5E007E6F0 CRC64;

```

Query Match 77.0%; Score 2627; DB 11; Length 649;
 Best Local Similarity 80.3%; Pred. No. 1e-181;
 Matches 530; Conservative 44; Mismatches 74; Indels 12; Gaps 7;

```

QY      1 MALTVSRSPGSGASTPVG-WDQAVQSRSLQRQSFAYLGAVLGLQDGGNDDAE 59
Db      1 MALTVSRSPGSGASTPVGPTQDRVVRRRRLQRQSFAYLGAVLGLQDGGSNVASE 60

```

```

QY      60 ASSEPTEKAPSEELHGDQDTEFGGQSOSPOKOEQRQHLHMVQLRPQDDIRLAAQLEA 119
Db      61 ADSEPMEEPSGEEQPTEDQTDKQGLQSPWKQ-VQKHHLHMLVBLRPQDDIRLAAQLEA 119
QY      120 RPPRLRLYLAVSTREGEGLSQDETVLLGVDFPDSSSPSCTGLVPLPMSDTQVYLDGDG 179
Db      120 RPPRLRLYLAVSTREGEGLSQDETVLLGVDFPDSSSPSCTGLVPLPMSDTQVYLDGDG 179
QY      180 GFSVTSGGQSRIFKPISIQTMATLOVLHQAACEALGSLVPGGSALTWASHYQERLNSE 239
Db      177 GFSVTSGGQSRIFKPVISIQTMATLOVLHQAACEVALGSLVPGGSALTWASHYQERLNSE 236
QY      240 QSCNWTAMADLESRLPPSAEPGSSSEQOEQMEQAIIRAEIMKVLVDVSDLESVTSKEIRQA 299
Db      237 QSCNWTAMADLESRLPPSAEPGSSSEQOEQMEQAIIRAEIMKVLVDVSDLESVTSKEIRQA 296
QY      300 LEHLGLPQQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQNRVTHIL 359
Db      297 LEHLGLPQQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQNRVTHIL 356
QY      360 MAREIDNFEYPERFTYHNVRWLWDEBSAQLLPHWKETHRFIEARAQGTHTVLVHCKMGVSR 419
Db      357 MAREIDNFEYPERFTYHNVRWLWDEBSAQLLPHWKETHRFIEARAQGTHTVLVHCKMGVSR 416
QY      420 SAATVLAYAMKQYECSEALRHVQELRPIARPNGFLRQLQIYQIGILTASQSHWEQKV 479
Db      417 SAATVLAYAMKQYECSEALRHVQELRPIARPNGFLRQLQIYQIGILTASQSHWEQKV 476
QY      480 VGVSPPEHPAPEVSTPFPPLPPEPEGGGEEKVGMESQAAPKEEPGRPRINLRGVMS 539
Db      477 VGVSPPEHPAPEVSTPFPPLPPEPEGGGEEKVGMESQAAPKEEPGRPRINLRGVMS 536
QY      540 SISLLEPSLEESTSETSDMPEVFSHSSHEEPLQFPQLARTKGQGVDRGPQALKS 599
Db      537 SISLLEPS-EEESTPEAGLPEVF---SSDEPLHPSQLSRAGGQVRVKGFWPALKS 591
QY      600 QSVVTLOGSAVVANRTQAFQEQEQGQGGQGEPCISSTPRFRKVVRAQASVDSGEEGEA 659
Db      592 QSVVTLOGSAVVANRTQAFQEQEQGQGGQGEPCISSTPRFRKVVRAQASVDSGEEGEA 649

```

RESULT 3

```

Q8N9L8 PRELIMINARY; PRT; 513 AA.
ID Q8N9L8;
AC Q8N9L8;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein FLJ36907.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Cerebellum;
RA Tanigami A., Fujiwara T., Shibahara T., Goto Y., Hirao M., Shimizu F.,
RA Wakebe H., Ono T., Hishigaki H., Watanabe T., Ozaki K., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK094226; BAC04314.1;
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.

```

KW Hypothetical protein; Kinase.
SQ SEQUENCE 513 AA; 56909 MW; 7F1FDA403C229A09 CRC64;
Query Match 76.9%; Score 2623; DB 4; Length 513;
Best Local Similarity 99.0%; Pred. No. 1.4e-181;
Matches 506; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 151 FP--DSSSPSCTGLVLPPLMSDTQVYLLDGDGFSVTSGQSRIFKPISTQTMATLQVH 208
Db 3 FPLSPASSPSCTGLVLPPLMSDTQVYLLDGDGFSVTSGQSRIFKPISTQTMATLQVH 62
QY 209 QACEALGSGLVPGSALTWASHYQERLNSEQSCINEMWTAMADLESIRPPSAEPGSSSQ 268
Db 63 QACEALGSGLVPGSALTWASHYQERLNSEQSCINEMWTAMADLESIRPPSAEPGSSSQ 122
QY 269 EQMEQAIRAELMKVLDVSDLESVTSKEIRQALERLGLPLQQRDFIDNOMLLVAQRDR 328
Db 123 EQMEQAIRAELMKVLDVSDLESVTSKEIRQALERLGLPLQQRDFIDNOMLLVAQRDR 182
QY 329 ASRIFFHLYLGESEWNAANLEELQNRVTHILNMAEIDNFPYPERFTYHNVRIMDEESAQL 388
Db 183 ASRIFFHLYLGESEWNAANLEELQNRVTHILNMAEIDNFPYPERFTYHNVRIMDEESAQL 242
QY 389 LPHWKETHRFIEARAQGTIVLVHCKMGVSRSAATVLAAMKQYECGLEQALRHVQELRP 448
Db 243 LPHWKETHRFIEARAQGTIVLVHCKMGVSRSAATVLAAMKQYECGLEQALRHVQELRP 302
QY 449 IARPNEGFLRQLIYQILTASRQSHWEQKGVSPPEHPAPEVSTPEPPLPEPEGGG 508
Db 303 IARPNEGFLRQLIYQILTASRQSHWEQKGVSPPEHPAPEVSTPEPPLPEPEGGG 362
QY 509 EEKVVMEESSQAAPKEEPGRPRINLRGVMRSISLEPSELESTSETSDMPEVSSHES 568
Db 363 GEKVVMEESSQAAPKEEPGRPRINLRGVMRSISLEPSELESTSETSDMPEVSSHES 422
QY 569 SHEEPLOPFPOLARTKGQGVDRGPQPALKRSQSVVTLQGSAAVAVANRTOAFQEQEQGGQ 628
Db 423 SHEEPLOPFPOLARTKGQGVDRGPQPALKRSQSVVTLQGSAAVAVANRTOAFQEQEQGGQ 482
QY 629 GQGEPCISSTPRFRKVVRAQSVHDSGEEGEA 659
Db 483 GQGEPCISSTPRFRKVVRAQSVHDSGEEGEA 513
RESULT 4
Q8WYLO PRELIMINARY; PRT; 471 AA.
AC Q8WYLO;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE HSSH-3.
GN HSSH-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21822082; PubMed=11832213;
RA Niwa R., Nagata-Ohashi K., Takeichi M., Mizuno K., Uemura T.;
RT "Control of Actin Reorganization by Slingshot, a Family of
RT Phosphatases that Dephosphorylate Adf/Cofilin.";
RL Cell 108:233-246(2002).
DR EMBL; AB072360; BAB84119.3; -.
DR HSSP; Q16828; IMKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PSS0056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PSS0054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 471 AA; 52741 MW; CE4E7F46FB562B2D CRC64;

Query Match 71.0%; Score 2422; DB 4; Length 471;
Best Local Similarity 99.8%; Pred. No. 4.4e-167;
Matches 468; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MALTVSRSPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGDNDDAEA 60
Db 1 MALTVSRSPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGDNDDAEA 60
QY 61 SSEPTEKAPSEELHGDQTDGQGSQSPQKEBQRQHLHMVQLLRPQDDIRLAQLEAP 120
Db 61 SSEPTEKAPSEELHGDQTDGQGSQSPQKEBQRQHLHMVQLLRPQDDIRLAQLEAP 120
QY 121 RPPRLRYLLVSTREGEGLSODETVLLGVDPPDSSSPSCTGLVLPPLMSDTQVYLLDGDG 180
Db 121 RPPRLRYLLVSTREGEGLSODETVLLGVDPPDSSSPSCTGLVLPPLMSDTQVYLLDGDG 180
QY 181 FSVTSGGQSRIFKPISTQTMATLQVLAQCEAALGSGLVPGSALTWASHYQERLNSEQ 240
Db 181 FSVTSGGQSRIFKPISTQTMATLQVLAQCEAALGSGLVPGSALTWASHYQERLNSEQ 240
QY 241 SCLENTWAMADLESRLRPPSAEPGSSSEQEQMEQAIRAELMKVLDVSDLESVTSKEIRQAL 300
Db 241 SCLENTWAMADLESRLRPPSAEPGSSSEQEQMEQAIRAELMKVLDVSDLESVTSKEIRQAL 300
QY 301 ELRLGLPLQQRDFIDNOMLLVAQRDRASRIFFHLYLGESEWNAANLEELQNRVTHILN 360
Db 301 ELRLGLPLQQRDFIDNOMLLVAQRDRASRIFFHLYLGESEWNAANLEELQNRVTHILN 360
QY 361 MAREIDNFPYPERFTYHNVRIMDEESAQLLPHWKETHRFIEARAQGTIVLVHCKMGVSR 420
Db 361 MAREIDNFPYPERFTYHNVRIMDEESAQLLPHWKETHRFIEARAQGTIVLVHCKMGVSR 420
QY 421 AATVLAAMKQYECGLEQALRHVQELRPIARPNEGFLRQLIYQILTA 469
Db 421 AATVLAAMKQYECGLEQALRHVQELRPIARPNEGFLRQLIYQILTA 469
RESULT 5
Q9NWZ7 PRELIMINARY; PRT; 394 AA.
AC Q9NWZ7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein FLJ20515.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Matanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isegai T., Sugano S.;
RT "MEDO human cDNA sequencing project.";
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000522; BAA91228.1; -.
DR KW
KW Hypothetical protein.
SQ SEQUENCE 394 AA; 42724 MW; F87C7F5A29048AAE CRC64;
Query Match 55.2%; Score 1884.5; DB 4; Length 394;
Best Local Similarity 59.8%; Pred. No. 2.9e-128;
Matches 394; Conservative 0; Mismatches 0; Indels 265; Gaps 1;
QY 1 MALTVSRSPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGDNDDAEA 60
Db 1 MALTVSRSPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGDNDDAEA 60
QY 61 SSEPTEKAPSEELHGDQTDGQGSQSPQKEBQRQHLHMVQLLRPQDDIRLAQLEAP 120
Db 61 SSEPTEKAPSEELHGDQTDGQGSQSPQKEBQRQHLHMVQLLRPQDDIRLAQLEAP 120

QY 121 RPPRLKLLVSTREGESQDETVLLGVDFPDSSSPSCTLGLVPLMSDTQVYLDGDDG 180
DB 121 RPPRLKLLVSTREGESQDETVLLGVDFPDSSSPSCTLGLVPLMSDTQVYLDGDDG 180
QY 181 FSVTSGGQSRIKFKPISIQTMWATLQVLHQAACEALGSLVPGSALTWASHYQERLNSQ 240
DB 181 FSVTSGGQSRIKFKPISIQTMWATLQVLHQAACEALGSLVPGSALTWASHYQERLNSQ 240
QY 241 SCINWETAMADLESIRPPSABPGSSSQEQMEQAIRAELMKVLDVSDIESVTSKEIRQAL 300
DB 241 SCINWETAMADLESIRPPSABPGSSSQEQMEQAIRAELMKVLDVSDIESVTSKEIRQAL 300
QY 301 ELRLGLPQYRDFIDNQMLLIVAQRDRASRIFPHLYLGSEWNAANLELQRNRYTHILN 360
DB 283 ----- 282
QY 361 MAREIDNFFPERFTYHNVRLEWDEESAQLLPHWKETHRFIEARAQCTHVLVHCKMGVSRS 420
DB 283 ----- 282
QY 421 AATVLAYAMKQYECSEQLRHVQELRPIARPNGFLRLQIYQGITLARSQSHWEQKV 480
DB 283 ----- 282
QY 481 GGVSPHEHPAPEVSTPFPPLPEPEGGEKXVGMESQAAPKEPGRPRINLRGVMRS 540
DB 283 ----- 282
QY 541 ISLLEPSLEESTSETSDMEPEVSSHESHSEEPLOPFPOLARTKGQGVDRGPQALKSR 600
DB 283 -----LEESTSETSDMEPEVSSHESHSEEPLOPFPOLARTKGQGVDRGPQALKSR 335
QY 601 QSVVTLQGSAAVAVNRTQAFQEQEQGQGGQGGEPCTISSTPRFRKVVRAQSVHDSGEEGEA 659
DB 336 QSVVTLQGSAAVAVNRTQAFQEQEQGQGGQGGEPCTISSTPRFRKVVRAQSVHDSGEEGEA 394

RESULT 6

Q9BQ20 PRELIMINARY; PRT; 299 AA.
ID Q9BQ20
AC Q9BQ20;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Similar to hypothetical protein FLJ10928.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004210; AAH04210.1; -.
DR EMBL; BC004176; AAH04176.1; -.
DR HSSP; P51452; 1VHR.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 299 AA; 33289 MW; B0C4BPEC16A23380 CRC64;

Query Match 45.8%; Score 1564; DB 4; Length 299;
Best Local Similarity 100.0%; Pred. No. 3.1e-105;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 MAREIDNFFPERFTYHNVRLEWDEESAQLLPHWKETHRFIEARAQCTHVLVHCKMGVSRS 420
DB 1 MAREIDNFFPERFTYHNVRLEWDEESAQLLPHWKETHRFIEARAQCTHVLVHCKMGVSRS 60
QY 421 AATVLAYAMKQYECSEQLRHVQELRPIARPNGFLRLQIYQGITLARSQSHWEQKV 480

DB 61 AATVLAYAMKQYECSEQLRHVQELRPIARPNGFLRLQIYQGITLARSQSHWEQKV 120
QY 481 GGVSPHEHPAPEVSTPFPPLPEPEGGEKXVGMESQAAPKEPGRPRINLRGVMRS 540
DB 121 GGVSPHEHPAPEVSTPFPPLPEPEGGEKXVGMESQAAPKEPGRPRINLRGVMRS 180
QY 541 ISLLEPSLEESTSETSDMEPEVSSHESHSEEPLOPFPOLARTKGQGVDRGPQALKSR 600
DB 181 ISLLEPSLEESTSETSDMEPEVSSHESHSEEPLOPFPOLARTKGQGVDRGPQALKSR 240
QY 601 QSVVTLQGSAAVAVNRTQAFQEQEQGQGGQGGEPCTISSTPRFRKVVRAQSVHDSGEEGEA 659
DB 241 QSVVTLQGSAAVAVNRTQAFQEQEQGQGGQGGEPCTISSTPRFRKVVRAQSVHDSGEEGEA 299

RESULT 7

Q8WYL4 PRELIMINARY; PRT; 692 AA.
ID Q8WYL4
AC Q8WYL4;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE HSSH-1S.
GN HSSH-1S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21822082; PubMed=11832213;
RA Niwa R., Nagata-Ohashi K., Takeichi M., Mizuno K., Uemura T.;
RT "Control of Actin Reorganization by Slingshot, a Family of
Phosphatases that Dephosphorylate ADF/Cofilin.";
RL Cell 108:233-246(2002).
DR EMBL; AB072356; BAB84115.1; -.
DR HSSP; Q16828; 1MKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 692 AA; 77430 MW; 9DC1FF2FC8984384 CRC64;

Query Match 34.3%; Score 1172; DB 4; Length 692;
Best Local Similarity 42.3%; Pred. No. 2.4e-76;
Matches 269; Conservative 95; Mismatches 170; Indels 102; Gaps 18;

QY 1 MALVTVSRSPPGSGASTPVGPWD---QAVQRSRLQROSFVLRGAVLGDGDNDD 56
DB 1 MALVTLSRSPTPSAASSASNSLEAGSEDRKNTLSSEFFWVGAAFLQGG----- 55
QY 57 AAEASSEPTKAPSEELHGDQTFQGS-QSPQKE-EQROHLHLMVQLRPDDIRLA 114
DB 56 -----SSPQ-----GQSLQHPKHAGDLPQHLQVMINILRCEDRIKLA 94
QY 115 AQLAPRPPRLKLLV--STREGESQDETVLLGVDFPDSSSPSCTLGLVPLMSDTQ 172
DB 95 VRLESAMADRVRYVVVSSGRQ----DTEENITLGVDFSSKESKSCITGMVLRMSDTK 150
QY 173 VYLDGDDGFSVTSGGQSRIKFKPISIQTMWATLQVLHQAACEALGSLVPGSALTWASHY 232
DB 151 IHLGDGDFSVSTAGRMHIFKPVSVQAMMSALQVLHKACEVAFRHNYPFGVALIATYY 210
QY 233 QERLNSQSCINWETAMADLESIRP-PSABPGSSSQEQMEQAIRAELMKVLDVSDLES 290
DB 211 ESCISSEQSCINWETAMADLESIRPSPALFVDKPTGERTERLLIKAKLSIMSGDLEN 270
QY 291 VTSKEIRQALERLGLPLQYRDFIDNQMLLIVAQRDRASRIFPHLYLGSEWNAANLEEL 350
DB 271 VTSKEIRNELEKQMCNKLKELKEFIDNEMILLILGQMDKPSLIFDHLILGSEWNASNLEEL 330

[illegible]

RESULT 8	
Q8WYL5	
ID	Q8WYL5
AC	Q8WYL5;
DT	01-MAR-2002 (Tremblrel, 20, Created)
DT	01-MAR-2002 (Tremblrel, 20, last sequence update)
DT	01-MAR-2003 (Tremblrel, 23, last annotation update)

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21822082; PubMed=11832213;
 RA Niwa R., Nagata-Chashi K., Takeichi M., Mizuno K., Uemura T.;
 RT "Control of Actin Reorganization by Slingshot, a Family of
 phosphatases that Dephosphorylate ADF/Cofilin.";
 RL Cell 108:233-246(2002).
 DR EMBL; AB072355; BAB84114.1; -.
 DR HSSP; Q16828; LMKP.
 DR InterPro; IPR000340; DS_phosphatase.
 DR InterPro; IPR000387; TYR_phosphatase.
 DR Pfam; PF00782; DSPC; 1.
 DR SMART; SM00195; DSPC; 1.
 DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
 DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
 SQ SEQUENCE 1049 AA; 115534 MW; 060C0F7C5E17889C CRC64;

Query Match	34.3%;	Score 1172;	DB 4;	Length 1049;
Best Local Similarity	42.3%;	Pred. No. 4.4e-76;		
Matches 269;	Conservative 95;	Mismatches 170;	Indels 102;	Gaps 18;

```

QY 1 MALVYTSRSPPGSGASTPVGPWD-----QAVQRRSRLQRRGFAVLRGAVLGLQDGGDND 56
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MALVTLORSPTPSAASSASNSLEHAGSEEDRKLNLSISEFFMWKGALFLQGG----- 55
QY 57 AAASSEPTTEKAPSEELHGDQTFGQGS-QSPQKQE-BQRHLHLMVQLLRPQDDIRLA 114
   |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 56 -----SSPQ-----GQSLQHPKHAGDLPQHLQVMINLRCEDRIKLA 94
QY 115 AQLEAPRPPRLRYLLV--STREGEGLSQDETVLLGVDFPDSSSPSCTLGLVPLMSDTQ 172
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 95 VRLESAMADRVRYMVVVYSSGRQ----DTEENILLVDFSSKESKSCSTIGVLRIMSDTK 150
QY 173 VYLDGDCGFSVTSGGQSRIFKPISIQTWMATLOVLHQAACAAALGSLVPGGSALTWASHY 232
   :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 151 IHLDDGCGFSVSTAGRMTIFKPVSVQAMMSALQVLHAKACEVARRHNYFPGGVALIMATYY 210
QY 233 QERLNSQSCLNWETAMADLESIRP--PSAEPGGSSSEQEQMEQAIRAELWKVLVDLES 290

```

[illegible]

RESULT 9
Q8N9A7
ID Q8N9A7 PRELIMINARY; PRT; 703 AA.

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Hypothetical protein FLJ38102.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.,
RT "NEDO human cDNA sequencing project."
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK035421; BAC04546.1; -.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hypothetical protein.
SQ SEQUENCE 703 AA; 78978 MW; 3ACFAB4B97C8DF7A CRC64;

Query Match	33.3%;	Score 1136;	DB 4;	Length 703;
Best Local Similarity	42.9%;	Pred. No. 1e-73;		
Matches 256;	Conservative 91;	Mismatches 152;	Indels 98;	Gaps 17;

```
QY      36 Q$FAVLKGAVLGLQDGGDNDDAAEASSEPTKEAPSEEBELHGDQTDFGGS-QSPÖKOE-E   93
       :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      51 ESFFWVGALLFLQG-----SSPQ-----GQSLQHPRKHAGD   84

QY      94 QRCHILMVQLRPQDDRILAQLEA PRPPRLRYLLV--STREGGLSQDETVLLGVDF   151
       :|::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      85 LPÖHLQWINLRLRCEDRIKLAVRESAMADRVRRVMVVVSSGRQ---DTENILLGVDF   140
```

```
QY 152 PDSSSPCTGLVLPFLMSDTQVYLDGGGFSVTSGGQSRIFKPISTQTMATLOVLHQAAC 211
Db 141 SSKESKSTCTIGWLRMLWSDTKIHLDDGGFSVSTAGRMHIFKPVSVQAMMSALQVLHKAAC 200
QY 212 EALGSGLVPGGSALTWASHYQERLNSQSCINWMTAMADLESIRP--PSAEPGGSSSEOE 269
Db 201 EVARRHNYFPGGVALIMATYYESCISSEQSCINWMTAMADLESIRP--PSAEPGGSSSEOE 260
QY 270 QMEQALRAELMKVLDVSDLESVTSKEIRQALRLGLPLQYRDFIDNOMLLVAQRDRA 329
Db 261 RTEKLAKAKRSIMSDLENVTSKEIRNELKQMCNCKELKEPIDNEMLLIGQMDKP 320
QY 330 SRIFPHLYLGSEWNAANLELQNRVTHILMAREIDNFPERFTYHNRVRLWDESSAQL 389
Db 321 SLIFDHLVYLGSEWNASNLEELQSGVDYILNVTREIDNFPGLFAYHNIRVDEETDIL 380
QY 390 PHMKETHRFIEARAOQTHVLVHCKMGVSRSAATVLAAMQYECSEALRHVQELRPI 449
Db 381 AHMNEAYHFINKAKRHSKCLVHCKMGVSRSAATVLAAMQYECSEALRHVQELRPI 440
QY 450 ARNPGFLRQLOIYQGITLTSRQSH--VWEQKVG--VSPREHPA-----PE 492
Db 441 TRPNAGFMRQLESEYEGILDASKORHKNLWQOTDSSLQYVDDPAGPGDFLPETPDGTP 500
QY 493 VSTPF-----PPLP-----PEPEGGEKVGMEB-----SQAP 522
Db 501 SOLPFLDDAQPGLGPPPLCCFRRLSDPLPSPEDETG--SLVHLEDPERALLLEAP 557
QY 523 KEE---PGRPRINLRGVMKRSISLEPSL--ELESTSETMDPEVFSHSSHEPL 574
Db 558 PAEVHRPARQPO-----QSGGLCEKDVKKLEFGSPKRSGLLOVEETREEG 607
```

RESULT 10

```
Q8C241 PRELIMINARY; PRT; 464 AA.
AC Q8C241;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE Similar to HSSH-1L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK089308; BAC40835.1;
SQ SEQUENCE 464 AA; 52781 MW; 96B736009573DDBC CRC64;
```

Query Match 33.2%; Score 1133; DB 11; Length 464;
Best Local Similarity 48.1%; Pred. No. 9.3e-74;
Matches 232; Conservative 85; Mismatches 129; Indels 36; Gaps 7;

```
QY 1 MALVTVSRSPPGSGASTPVGPWDQAV---QRRSRLQRQSFVLRGAVLGLQDGDND 56
Db 1 MALVTIQRSPFPAASSASNSSELEAGSDEBRKLNLSSESFFMVKGAALFLQG----- 55
QY 57 AAASSEPTKAPSEELHGDQDTDFGQGS-QSPQKE-EQRQHLHMLVQLLRPODDIRLA 114
Db 56 -----NSPQ-----GQSLQHPKHAGDLPQHLQVMINILRCEDRIKLA 94
QY 115 AQLAPRPRLRYLLVSTREGESLQDETIVLGVDFPDSSSPCTGLVPLMSDTQVY 174
Db 95 VRLESVWTDVRYVMVVTYSGRQ--DTEENILGVDFSSSKESKSTCTIGWLRMLWSDTKI 152
```

```
QY 175 LDGGGFSVTSGGQSRIFKPISTQTMATLOVLHQAACEALGSGLVPGGSALTWASHYQ 234
Db 153 LDGGGFSVSTAGRMHIFKPVSVQAMMSALQVLHKAACEVARRHNYFPGGVALIMATY 212
QY 235 RLNSQSCINWMTAMADLESIRP--PSAEPGGSSSEOEQMEQALRAELMKVLDVSD 292
Db 213 CISSEQSCINWMTAMADLESIRP--PSAEPGGSSSEOEQMEQALRAELMKVLDVSD 272
QY 293 SKEIRQALRLGLPLQYRDFIDNOMLLVAQRDRAIRFPHLYLGSEWNAANLELQ 352
Db 273 SKEIRNELKQMCNCKELKEPIDNEMLLIGQMDKPSLIFDHLVYLGSEWNASNLEELQ 332
QY 353 NRVTHILMAREIDNFPERFTYHNRVRLWDESSAQLPHMKETHRFIEARAOQTHVLV 412
Db 333 SGVDYILNVTREIDNFPGLFAYHNIRVDEETDILAHMNEAYHFINKAKRHSKCLV 392
QY 413 CKMGVSRSAATVLAAMQYECSEALRHVQELRPIARNPGFLRQLOIYQGITLTSRQ 472
Db 393 CKMGVSRSAATVLAAMQYECSEALRHVQELRPIARNPGFLRQLOIYQGITLTSRQ 452
QY 473 SH 474
Db 453 RH 454
```

RESULT 11

```
Q8WYL2 PRELIMINARY; PRT; 449 AA.
AC Q8WYL2;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, last annotation update)
DE HSSH-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21822082; PubMed=11832213;
RA Niwa R., Nagata-Ohashi K., Takeichi M., Mizuno K., Yemura T.;
RT "Control of Actin Reorganization by Slingshot, a Family of
RL Cell 108:233-246 (2002).";
DR EMBL; AB072358; BAB84117.1;
DR HSSP; Q16828; IMKP.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 449 AA; 51492 MW; AFD156098A92A04D CRC64;
```

Query Match 31.9%; Score 1089.5; DB 4; Length 449;
Best Local Similarity 44.7%; Pred. No. 1.3e-70;
Matches 215; Conservative 97; Mismatches 124; Indels 45; Gaps 5;

```
QY 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSRLQR---QRAVLRGAVLGLQDGDND 57
Db 1 MALVTIQRSPFPAASSASNSSELEAGSDEBRKLNLSSESFFMVKGAALFLPR----- 53
QY 58 AAASSEPTKAPSEELHGDQDTDFGQGSQSPQKEE-----RQHLHMLVQLLRPODD 110
Db 54 -----GNGSSTPRISHRRNKHAGDLPQHLQVMINILRCEDRIKLA 89
QY 111 IRLAQLAPRPRLRYLLVSTREGESLQDETIVLGVDFPDSSSPCTGLVPLMSD 170
Db 90 IRLAVRLESTYQNRTRYVMVVTYSGRQ--DTEESIVLGMDFSSNDSTCTMGLVPLMSD 147
QY 171 TQVYLDGGGFSVTSGGQSRIFKPISTQTMATLOVLHQAACEALGSGLVPGGSALTWAS 230
```


Db 148 TLHLDGDFSVSTDRVHIFKPVSVQAMWSALQSLHAKACEVAPRANYYPSGLFTWVS 207
QY 231 HYGERLNSEQSLNWTAMADLESRLP--PSAEPGSSSEQEQMEQAIRAELMKVLDVSD 288
Db 208 YYESHINSQSSVNEWMAMQDVQSHRPDSPALFTDIPTERERTERLTKTLREIMQKDL 267
QY 289 ESVTSKEIRQALELRGLPLQCYRDFIDNQMLLVAQRDRASRIFPHLYLGSEWNAANLE 348
Db 268 ENITSKEIRTELEMQMCULREFKEFIDNEMIVILGQMSPTQIFEHVFLGSEWNASNLE 327
QY 349 ELQNRVTHILNMAREIDNFPERFTYHNVRMLDEESAQLLPHMKETHRFIEARAQGT 408
Db 328 DLQNRGVRYILNVTREIDNFPFGVFYHNIRVYDEATDILAYMNDTYKFKISKAKKGSK 387
QY 409 VLVHCKMGVSRSAATVLAAMKQYECLEQALRHVQELRPIARPNGFLRLQIYQGITL 468
Db 388 CLVHCKMGVSRSAATVLAAMKQYECLEQALRHVQELRPIARPNGFLRLQIYQGITL 447
QY 469 A 469
Db 448 A 448

RESULT 12

Q9NRY1 PRELIMINARY; PRT; 1045 AA.
ID Q9NRY1 Q9NRY1: Q9VC04;
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE MAP kinase phosphatase (MAP-kinase-phosphatase protein).
GN SSH OR CG6238.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Niwa R., Takeichi M., Uemura T.;
RT "a novel MAP kinase phosphatase (Drosophila).";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskearn D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Snue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
DR EMBL; AB036834; BAA89534.1; -
DR EMBL; AE003750; AAF56372.2; -
DR HSSP; Q16828; IMKP.
DR FlyBase; FBgn0029157; ssh.
DR InterPro; IPR000340; DS_phosphatase.
DR InterPro; IPR000387; TYR_phosphatase.
DR Pfam; PF00782; DSPC; 1.
DR SMART; SM00195; DSPC; 1.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.
KW Hydroxylase; Kinase.
SQ SEQUENCE 1045 AA; 114996 MW; 76620BB63EE58A8B CRC64;

Query Match 27.6%; Score 940.5; DB 5; Length 1045;
Best Local Similarity 34.6%; Pred. No. 2.5e-59;
Matches 230; Conservative 96; Mismatches 219; Indels 119; Gaps 11;
QY 1 MALVTYSRSPGSGA-STVPQVWDQAVQKRSRLQRRQSFVLRGAVLGLQDGDNDAAE 59
Db 1 MALVTQSRSPSVAGSCSNDSDESEDEGNSKGNDRSECFPAKGATLVL----- 49
QY 60 ASSEPTKAPSEELHGDQTFEGQSQSPQKQEQROHLHLMVOLRPPDIRLAQLEA 119
Db 50 ALKDIPPIQSERRLSTDSRSSNSTQS--NNSDIQLHLSQSMFYLLQREDTLKVAVKLES 107
QY 120 PRPRRLRYLVVS----- 132
Db 108 QRSNRTRYLVIASRSCCRSGTSDRRHRIRMRHSYKVGSGAGTKSSTPAVPTQRLSVE 167
QY 133 -----TREGEGLS--QDETVLGVDFPDSSSPCTGLVPLWSD 170
Db 168 QTATBASCKDKTADKENATAAGDNKNTSGMEESCLGIDCNERT---TIGLVPIAD 223
QY 171 TQVYLDGDFSVTSQSGQSRIFKPISIQTMWATQVLHQAQEAALGSGLVGSGALTWAS 230
Db 224 TTIHLDGDFGSVKVEKTHIFKPVSVQAMWSALQTLHKVSKKARENNEFYASGSHDWLS 283
QY 231 HYGERLNSEQSLNWTAMADLESRLPSPAEP--GSSSEQEQMEQAIRAELMKVLDVSD 288
Db 284 SYERRIESQSLNEMNAMADLESRRPSPDAIRNKPPEKETESVTKMLKAIMSVDL 343
QY 289 ESVTSKEIRQALELRGLPLQCYRDFIDNQMLLVAQRDRASRIFPHLYLGSEWNAANLE 348
Db 344 DEVTSKYIRGRLEILLDMGLGEYKSFIDAEMVLVIGOMDAPTKIFEHVILGSEWNASNLE 403
QY 349 ELQNRVTHILNMAREIDNFPERFTYHNVRMLDEESAQLLPHMKETHRFIEARAQGT 408
Db 404 ELQNRGVRYILNVTREIDNFPFGVFYHNIRVYDEATDILAYMNDTYKFKISKAKKGSK 463
QY 409 VLVHCKMGVSRSAATVLAAMKQYECLEQALRHVQELRPIARPNGFLRLQIYQGITL 468
Db 464 VLVHCKMGVSRSAATVLAAMKQYECLEQALRHVQELRPIARPNGFLRLQIYQGITL 523
QY 469 ASR-----QSHWEQYGVSPPEHPAPV-----STPEPLPPEEGGE 509
Db 524 AMKNKELQRSKSETNLKSTKDARLLPGSEPTPLQALNQAKSKSGTAGVTPD--GEE 580
QY 510 EKVNGMESQAAPKEEPGRPRINLRGVMRSISLLEPSLESTSETSDMPEVSSHSS 569
Db 581 EDGSRMRRSIAQKQ-----RRMVRSSSTSPKQTAVVTQKQSGSMENLTPERSV 632

QY 570 HEEP 573
DB 633 AEEP 636

RESULT 13
O8IMU8 PRELIMINARY; PRT; 1193 AA.
AC O8IMU8;
DT 01-MAR-2003 (TremBurel. 23, Created)
DT 01-MAR-2003 (TremBurel. 23, Last sequence update)
DT 01-MAR-2003 (TremBurel. 23, Last annotation update)
DE CG5238-PB.
GN SSH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; Pubmed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgeret, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galie R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,

RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

[5]
RP SEQUENCE FROM N.A.
RA Flybase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003750; AANI4027.1;
SQ SEQUENCE 1193 AA; 131521 MW; A1B9398A0443FF1B CRC64;

Query Match 26.9%; Score 918.5; DB 5; Length 1193;
Best Local Similarity 34.2%; Pred. No. 1.2e-57;
Matches 218; Conservative 96; Mismatches 192; Indels 131; Gaps 11;

QY 49 QDGDNDAAEASSEPTKA-----PSEELHGDQTFGQSQ 86
DB 20 RDGESEDD--EGNSKGNDRSECFAGKGTALVLAIKDIPLTQSERLSTDSRSTQ 77
QY 87 SPQKEBQROHLMVQLRPQDDIRLAQLEAPRPRLLVVS----- 132
DB 78 S--NNSDIQLHQSMPYLLQREDLTKMAVKLESQSRNRYLVIASRSCSGTSRRRH 135
QY 133 -----TREGEGLS 140
DB 136 RIMRHVSVKYGSAGTSTSPAVPTQRLSVEQTATEASSKCDKTADKENATAADNKN 195
QY 141 ---QDETVLGVDPDSSSPSCITGLVPLWMSDTQVYLDGDSFVTSGGQRIFFKPI 197
DB 196 TSGMEESCLIGIDCNERT---TIGLVPIADTTIHLDGDSFVTSVVEKTHFFKPSV 251
QY 198 QTMWATLOVLAQCEALGSGLVGGSALTWASHYQERLNSQSCINENWTAMADLES 257
DB 252 QAMWSALQTLHKVSKKARENNEFYASGPHDLSYSYERIESDQCLNENWAMDALES 311
QY 258 PSAEP--GGSSEQEQEQAIRAEIKVLDVDSLESVTSKEIRQALERLGLPQOYR 315
DB 312 PSPDAIRNKPPEKETESVIMKKAIMSVLDDEVTSKYIRGLLEILMDLGEYKSP 371
QY 316 DNOMLLVAQRDASRIFFPHLYLGSEWNAANLELQNRVTHILMAREIDNRYRFTY 375
DB 372 DAEMLVILQMDAPTIFEHVYLGSEWNASNLEELQNGVRHILNVTREIDNFPPTFEY 431
QY 376 HNVRLMDEESACQLPHWKETHRFTEARAQGTHTLVHCKGVSRSAATVLAAMKQY 435
DB 432 FNVRYVYDDEKTNLLKYWDTFRYITRAKAEGSKVLVHCKGVSRSAATVLAAMK 491
QY 436 LEQALRHVQELRPILARNPGFLROIYQILTAASR-----QSHWEQYGVGVSP 485
DB 492 FQQALEHVKKRSCIKPNKNFLNQLTYSGMLDAMKNKXELQRSKSETNLKSTDA 551
QY 486 BEHPAPEV-----STFPPLPEPEEGGEEKVGVMEESQAAPKEPGRPRINLRG 536
DB 552 GSEPTPLIQLANQAKSKSTGEAGVTPD---GSEEDGSRMRRSIAQKSO-----RRM 600
QY 537 VMSISLLEPSLESTSETSDMPVFSSHSSHEEP 573

Db 601 VRRSSSTSPKQTAVTVTKQSSQSMENLTPERSVAEEP 637

RESULT 14

Q9NV45

ID Q9NV45 PRELIMINARY; PRT; 141 AA.

AC Q9NV45; 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)

DE Hypothetical protein FLJ10928.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Ovarian carcinoma;

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,

RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,

RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,

RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,

RA Masuko Y., Kanehori K.;

RT "NEDO human cDNA sequencing project.;"

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK001790; BAA91913.1; -

DR HSSP; P51452; 1VHR.

DR InterPro; IPR000340; DS_phosphatase.

DR InterPro; IPR000387; TYR_phosphatase.

DR Pfam; PF00782; DSPC; 1.

DR SMART; SMO0195; DSPC; 1.

DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.

KW Hypothetical protein.

SQ SEQUENCE 141 AA; 16357 MW; 5B268790931CA5B6 CRC64;

QY

Query Match 17.0%; Score 581.5; DB 4; Length 141;

Best Local Similarity 89.0%; Pred. No. 1.5e-34;

Matches 113; Conservative 2; Mismatches 5; Indels 7; Gaps 1;

Db 350 LQNRVTHILN-----MAREIDNFFPERFTYHNVRRLTWDEESAQLLPHWKETHRFIEAA 402

Db 13 LSQHPSPFILNSPASHSPAREIDNFFPERFTYHNVRRLTWDEESAQLLPHWKETHRFIEAA 72

QY 403 RAQGTHTLVHCKMGVSRSAATVLAAYAMKQYECSLAQALRHVQELRPIARPNGFLRQLOI 462

Db 73 RAQGTHTLVHCKMGVSRSAATVLAAYAMKQYECSLAQALRHVQELRPIARPNGFLRQLOI 132

QY 463 YQGILTA 469

Db 133 YQGILTA 139

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

RT for large proteins in vitro.;

RL DNA Res. 7:65-73(2000).

DR EMBL; AB037719; BAA92536.1; -

DR HSSP; Q16828; 1MKP.

DR InterPro; IPR000340; DS_phosphatase.

DR InterPro; IPR000387; TYR_phosphatase.

DR Pfam; PF00782; DSPC; 1.

DR SMART; SMO0195; DSPC; 1.

DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS50054; TYR_PHOSPHATASE_DUAL; 1.

KW Hypothetical protein.

FT NON TER 1

SQ SEQUENCE 738 AA; 80784 MW; C92D75671A510F8A CRC64;

QY

Query Match 12.2%; Score 416.5; DB 4; Length 738;

Best Local Similarity 36.4%; Pred. No. 1.3e-21;

Matches 106; Conservative 37; Mismatches 85; Indels 63; Gaps 11;

Db 337 YLG-SENNANLLELQNRVTHILNMAREIDNFFPERFTYHNVRRLTWDEESAQLLPHWKET 395

Db 5 YLSWDRWTSPPSSIIFFISYDYIINVTREIDNFFPGLFAYHNIRVYDEETDLLAHWNEA 64

QY 396 HRFIEARAQGTHTLVHCKMGVSRSAATVLAAYAMKQYECSLAQALRHVQELRPIARPNG 455

Db 65 YHFINAKARNHNSKCLVHCKMGVSRSAATVLAAYAMKEFGWPLEKAYNYVKQKRSITRPNAG 124

QY 456 FLRQLOIYQGITLTASRQSH--VWEQKVG--VSEEHPA-----PEVSTPF- 497

Db 125 FMRQISEYEGLIDASKQRHNMKLWRQOTDSSLQQPVDDPAGPGDFLPETPDGTPESQLPFL 184

QY 498 -----PPLP-----PEPEGGEKVVGMEE-----SQAPKKE--- 525

Db 185 DDAAQPGGLGPPLPCCFRRLSDPLLPSPEDETG---SLVHLEDPEREALLLEAAPPAAVHR 241

QY 526 PGPFRPRINLRGVMSISLLEPSL--ELESTSETSDMPEVFSSHSHSHEEPL 574

Db 242 PARQPQ-----QGSGICEKDVKKTLFEGSPKGRSGSLQVEETREBEGL 285

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

Search completed: January 15, 2004, 06:49:23
Job time : 44 secs

FT Modified-site /note= "O-phosphorylated"
FT 266 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 286 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 292 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 484 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 542 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 552 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 564 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 568 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 649 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 653 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 65 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 254 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 395 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 638 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 369 /note= "O-phosphorylated"
FT Modified-site /note= "O-phosphorylated"
FT 615 /note= "O-phosphorylated"
FT Modified-site /note= "N-glycosylated"
XX
PN WO200120004-A2.
XX
PD 22-MAR-2001.
XX
PF 14-SEP-2000; 2000WO-US25515.
XX
PR 15-SEP-1999; 99US-0154141.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Tang YT, Bandman O, Hillman JL, Baughn MR, Azimzai Y;
PI Lu DAM;
XX
DR WPI; 2001-244811/25.
DR N-PSDB; AAF30485.
XX
PT Novel human protein phosphatase and kinase proteins for diagnosis,
PT treatment and prevention of gastrointestinal, immune system,
PT neurological and cell proliferative disorders -
XX
PS Claim 1; Page 93-94; 103pp; English.
XX
CC The present sequence is that of novel human protein phosphatase and
CC kinase protein PPHKP-10, as predicted from Incyte Clone ID No.
CC 5039718CB1 (see AAF30485). Tissues that express PPHKP-10 (as a
CC fraction of total tissues expressing PPHKP-10) include reproductive
CC (0.343), gastrointestinal (0.194) and haematopoietic or immune
CC (0.134). Diseases or conditions associated with tissues expressing
CC PPHKP-10 (as a fraction of total tissues expressing PPHKP-10)
CC include cancer (0.552), inflammation or trauma (0.314) or cell
CC proliferation (0.090). The encoded protein shows homology to
CC Drosophila melanogaster Map kinase phosphatase. The invention
CC provides human PPHKP-1 to -11 polypeptides (see AAB20322-32) and
CC polynucleotides (see AAF30476-86). It also provides expression
CC vectors, host cells, antibodies, agonists and antagonists, as well
CC as methods for diagnosing, treating or preventing disorders
CC associated with expression of PPHKP, including gastrointestinal

CC disorders, immune system disorders, neurological disorders and cell
CC proliferative disorders, including cancer.
XX
SQ Sequence 659 AA;
Query Match 100.0%; Score 3412; DB 22; Length 659;
Best Local Similarity 100.0%; Pred. No. 1e-281;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALTVSRSPGSGASTPVGPMQAVQRRRLQRRQSFVLRGAVLGLQDGDNDPAEA 60
DB 1 MALTVSRSPGSGASTPVGPMQAVQRRRLQRRQSFVLRGAVLGLQDGDNDPAEA 60
QY 61 SSEPTKAPSEELHGDQTDGQSGSPCKQEEORHLHLMVQLRPQDDIRLAQLEAP 120
DB 61 SSEPTKAPSEELHGDQTDGQSGSPCKQEEORHLHLMVQLRPQDDIRLAQLEAP 120
QY 121 RPRRLRYLLVSTREGELSDQDQTVLLGVDFPDSSSPSCITGLVPLWSDTQVYLDGCG 180
DB 121 RPRRLRYLLVSTREGELSDQDQTVLLGVDFPDSSSPSCITGLVPLWSDTQVYLDGCG 180
QY 181 FSVTSGGQSRIFKPISTQTMWATLQVLRQCEALGSGLVPGSALTWASHYQERLNEQ 240
DB 181 FSVTSGGQSRIFKPISTQTMWATLQVLRQCEALGSGLVPGSALTWASHYQERLNEQ 240
QY 241 SCLENTAMADLESLEPPSAEPGSSSEQEQEAIKRLMKVLDVSDLESVTSKEIRQAL 300
DB 241 SCLENTAMADLESLEPPSAEPGSSSEQEQEAIKRLMKVLDVSDLESVTSKEIRQAL 300
QY 301 ELRLGLPLQYRDFINQMLLLVAQRDRASRIFFHLYLGSEWNAANLELQNRVTHILN 360
DB 301 ELRLGLPLQYRDFINQMLLLVAQRDRASRIFFHLYLGSEWNAANLELQNRVTHILN 360
QY 361 MAREIDNFFPERFTYHNVRLWDEESAQLLPHWKETHRFIEAARQGTHTLVHCKMGYSRS 420
DB 361 MAREIDNFFPERFTYHNVRLWDEESAQLLPHWKETHRFIEAARQGTHTLVHCKMGYSRS 420
QY 421 AATVLAAYAMKQYECLEQALRHVQELRPIARPNGFLRQLQIYQILITASQSHWEQKV 480
DB 421 AATVLAAYAMKQYECLEQALRHVQELRPIARPNGFLRQLQIYQILITASQSHWEQKV 480
QY 481 GGVSPHEHPAPEVSTPPPLPPEPEGGEKVVMEESQAAPKEGPRPRINLRGVMS 540
DB 481 GGVSPHEHPAPEVSTPPPLPPEPEGGEKVVMEESQAAPKEGPRPRINLRGVMS 540
QY 541 ISLEPSLELESTSETSDMEVFSSHESHEBPLQPPOLARTGQGVDRGFQPALKSR 600
DB 541 ISLEPSLELESTSETSDMEVFSSHESHEBPLQPPOLARTGQGVDRGFQPALKSR 600
QY 601 QSVYTLQGSAAVVAARQAFQEQEQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGG 659
DB 601 QSVYTLQGSAAVVAARQAFQEQEQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGG 659
RESULT 2
ID AAE22729 standard; Protein; 659 AA.
XX
AC AAE22729;
XX
DT 09-AUG-2002 (first entry)
XX
DE Human dual-specificity phosphatase 15 (DSP-15) protein.
XX
KW Human; dual-specificity phosphatase 15; DSP-15; antiallergic; cytosolic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW metabolic disease; allergy; screening; chromosome 11q.
XX
OS Homo sapiens.
XX
PN WO200224740-A2.

XX 28-MAR-2002.
PD
XX
PF 19-SEP-2001; 2001WO-US29406.
XX
PR 19-SEP-2000; 2000US-233833P.
PR 18-SEP-2001; 2001US-0955732.
XX
PA (CEPT-) CEPTYR INC.
XX
PI Luche RM, Wei B;
XX
DR WPI; 2002-394127/42.
DR N-PSDB; AAD36061.
XX
XX
PT New dual-specificity phosphatase 15 polypeptide and polynucleotides,
PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
PT diseases -
XX
XX
PS Claim 1; Fig 2; 91pp; English.
XX
CC The invention relates to a new isolated dual-specificity phosphatase 15
CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
CC residues. DSP-15 polypeptides may be used to identify agents that
CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
CC polypeptides, modulating agents, and/or polynucleotides encoding the
CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
CC abnormal cell growth, abnormal cell proliferation and cell cycle
CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
CC assays for modulators of enzyme activity and/or substrate binding. The
CC present sequence is human DSP-15 protein. Human DSP-15 gene is located on
CC chromosome 11q.
XX
SQ Sequence 659 AA;

Query Match 100.0%; Score 3412; DB 23; Length 659;
Best Local Similarity 100.0%; Pred. No. 1e-281;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALVTVSRSPPGSGASTPVGPMQAVQRRSRLQRRSFAVLRGAVLGLDGGDNDADA 60
Db 1 MALVTVSRSPPGSGASTPVGPMQAVQRRSRLQRRSFAVLRGAVLGLDGGDNDADA 60

QY 61 SSEPTKAPSEELHGDQTFDGGGSSQSPQKEQGRHLHMLVQLRPDDIRLAQL 120
Db 61 SSEPTKAPSEELHGDQTFDGGGSSQSPQKEQGRHLHMLVQLRPDDIRLAQL 120

QY 121 RPPRLRYLLVSTREGGLSGDETVLLGVDFPDSSSPSCTLGLVPLMSDTQVYLDG 180
Db 121 RPPRLRYLLVSTREGGLSGDETVLLGVDFPDSSSPSCTLGLVPLMSDTQVYLDG 180

QY 181 FSVTSGGSRIFKPISIQTWATLQVLHQAACEALGSLVPGGSALTWASHYQERL 240
Db 181 FSVTSGGSRIFKPISIQTWATLQVLHQAACEALGSLVPGGSALTWASHYQERL 240

QY 241 SCINWNTAMADLESRLPPSAEFGSSSQEQMEQAIRAELMKVLDVSDLESVTSKEIR 300
Db 241 SCINWNTAMADLESRLPPSAEFGSSSQEQMEQAIRAELMKVLDVSDLESVTSKEIR 300

QY 301 ELRLGLPQQYRDFIDNOMLLVAQRDRASRIFFHLVLSGSEWNAANLEIQRNVT 360
Db 301 ELRLGLPQQYRDFIDNOMLLVAQRDRASRIFFHLVLSGSEWNAANLEIQRNVT 360

QY 361 MAREIDNYPFRFTYHNRLWDEESAQLLPHWKETHRFIEARAQGTIVLVHCKMGVSR 420
Db 361 MAREIDNYPFRFTYHNRLWDEESAQLLPHWKETHRFIEARAQGTIVLVHCKMGVSR 420

QY 421 AATVLAYAMKQYECLEQALRHVQELRPIARBPNPGLRQLQIYQGITLTSRSHVMEQKV 480
Db 421 AATVLAYAMKQYECLEQALRHVQELRPIARBPNPGLRQLQIYQGITLTSRSHVMEQKV 480

QY 481 GGVSPPEHPAPEVSTPFPPLPPEPEGGEKVVGMESQAAPKEPGRPRINLRGVMS 540
Db 481 GGVSPPEHPAPEVSTPFPPLPPEPEGGEKVVGMESQAAPKEPGRPRINLRGVMS 540

QY 541 ISLLEPSLESTSETSDMPEVFSHSSHEEPLOPPQALRTKGGQVDRGPQALRSR 600
Db 541 ISLLEPSLESTSETSDMPEVFSHSSHEEPLOPPQALRTKGGQVDRGPQALRSR 600

QY 601 QSVVTLQGSAAVYANRTQAFQEQEQGQGGGEPCCISSTPRRKVVROASVHDSGEGEA 659
Db 601 QSVVTLQGSAAVYANRTQAFQEQEQGQGGGEPCCISSTPRRKVVROASVHDSGEGEA 659

RESULT 3
ABP51654
ID ABP51654 standard; Protein; 659 AA.
XX
AC ABP51654;
XX
DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 2 protein SEQ ID NO:5.
XX
KW Human; phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11.
XX
OS Homo sapiens.
XX
FN WO200242436-A2.
XX
PD 30-MAY-2002.
XX
PF 07-NOV-2001; 2001WO-US42995.
XX
PR 20-NOV-2000; 2000US-0715177.
PR 18-JAN-2001; 2001US-0761640.
XX
XX
PA (PEKE) PE CORP NY.
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-575237/61.
DR N-PSDB; ABQ73250.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies -
XX
PS Claim 1; Fig 2A; 85pp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC represent human MAP kinase phosphatase splice form 2 from the present
CC invention.
XX
SQ Sequence 659 AA;

Query Match 99.8%; Score 3402; DB 23; Length 659;
Best Local Similarity 99.8%; Pred. No. 7.2e-281;
Matches 658; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MALVTVSRSPGSGASTPVGPWMDQVQRSLQRQRQSFAYLRGAVLGLQDGGDNDAAEA	60
Db	1	MALVTVSRSPGSGASTPVGPWMDQVQRSLQRQRQSFAYLRGAVLGLQDGGDNDAAEA	60
QY	61	SSEPTKAPSEELHGDQTFDGGSGSPQKQEEQRQHLHWQLLRPODDIRLAQLEAP	120
Db	61	SSEPTKAPSEELHGDQTFDGGSGSPQKQEEQRQHLHWQLLRPODDIRLAQLEAP	120
QY	121	RPRRLRYLLVSTREGEGLSQDETLLGVDFPDSSPSCTIGIVLPLMSDTQVYLDGCG	180
Db	121	RPRRLRYLLVSTREGEGLSQDETLLGVDFPDSSPSCTIGIVLPLMSDTQVYLDGCG	180
QY	181	FSVTSGGQSRIFKPISTQTMWATLVQACEALGSLVPGSALTWASHYQERLNSQ	240
Db	181	FSVTSGGQSRIFKPISTQTMWATLVQACEALGSLVPGSALTWASHYQERLNSQ	240
QY	241	SCLNWTAMADLESRLRPPSAEPGSSSQEOMEQAIRAELMKYLVDSLESVTSKEIRQAL	300
Db	241	SCLNWTAMADLESRLRPPSAEPGSSSQEOMEQAIRAELMKYLVDSLESVTSKEIRQAL	300
QY	301	ELRLGLPLQGYRDFIDNQMLLVQAQRDASRIFFPHLYLGSENNANLEELQRNRVTHLN	360
Db	301	ELRLGLPLQGYRDFIDNQMLLVQAQRDASRIFFPHLYLGSENNANLEELQRNRVTHLN	360
QY	361	MAREIDNFPERFTYHNVRWLDEESAQLPHWKETHRFIEARAQSTHVLVHCKMGVRS	420
Db	361	MAREIDNFPERFTYHNVRWLDEESAQLPHWKETHRFIEARAQSTHVLVHCKMGVRS	420
QY	421	AATVLAYAMKQYECSLFQALRHVQELRPIARPNGFLRQLIYQGITLTSRQSHWEQY	480
Db	421	AATVLAYAMKQYECSLFQALRHVQELRPIARPNGFLRQLIYQGITLTSRQSHWEQY	480
QY	481	GGVSPPEHPAPEVSTPFPPLPPEPEGGEKVVGMESQAAPKEHPGPRPRINLRGVRS	540
Db	481	GGVSPPEHPAPEVSTPFPPLPPEPEGGEKVVGMESQAAPKEHPGPRPRINLRGVRS	540
QY	541	ISLLEPSLELESTSETSDMPEVFSSHESHEEPLOFPQRLARTKGQOVDRGPQALKS	600
Db	541	ISLLEPSLELESTSETSDMPEVFSSHESHEEPLOFPQRLARTKGQOVDRGPQALKS	600
QY	601	QSVVTLQGSAAVVANRTQAFQEQEQGGQGGGEPICISSTPRFRKVVROASVHDSGEEGEA	659
Db	601	QSVVTLQGSAAVVANRTQAFQEQEQGGQGGGEPICISSTPRFRKVVROASVHDSGEEGEA	659
RESULT 4			
ID	ABB07845	standard; Protein; 779 AA.	
XX	AC		
XX	ABB07845;		
DT	03-JUL-2002	(first entry)	
XX			
DE		Human MAP kinase phosphatase-like enzyme.	
KW		Mitogen activated protein; MAP; MAP kinase phosphatase-like enzyme;	
KW		antidiabetic; anorectic; cytostatic; cardiant; human;	
KW		antiparkinsonian; cerebroprotective; neuroprotective; nootropic;	
KW		neuroleptic; anticonvulsant; anti-HIV; antiarrhythmic; hypotensive;	
KW		antiallergic; dermatological; vulnerary; gene therapy.	
XX			
OS		Homo sapiens.	
XX			
FH			
FT	Key	Location/Qualifiers	
FT	Misc-difference 53	/note= "the encoding codon is not indicated in the	
FT		corresponding DNA sequence"	
FT	Misc-difference 62	/note= "the encoding codon is not indicated in the	
FT		corresponding DNA sequence"	
FT	Misc-difference 141	/note= "the encoding codon is not indicated in the	
FT		corresponding DNA sequence"	

FT	Misc-difference	403	/note=	"the encoding codon is not indicated in the
FT				corresponding DNA sequence"
FT	Misc-difference	701	/note=	"the encoding codon is not indicated in the
FT				corresponding DNA sequence; Xaa can be any
FT				amino acid"
XX				
PN	WO200220732-A2.			
XX				
PD	14-MAR-2002.			
XX				
PF	27-AUG-2001; 2001WO-EP09848.			
XX				
PR	07-SEP-2000; 2000US-230709P.			
XX				
PA	(FARB) BAYER AG.			
XX				
PI	Liou J;			
XX				
DR	WPI; 2002-339802/37.			
DR	N-PSDB; ABL40805.			
XX				
PT	New human mitogen activated protein kinase phosphatase-like enzyme			
PT	polypeptide, regulators of which are useful for preventing, treating			
PT	allergies including asthma, diabetes, obesity, cancer and			
PT	cardiovascular diseases			
XX				
XX				
XX	Claim 18; Fig 14; 134pp; English.			
XX				
CC	The invention relates to a purified human mitogen activated protein (MAP)			
CC	kinase phosphatase-like enzyme polypeptide. The enzyme can be expressed			
CC	by standard recombinant methodology. The MAP kinase phosphatase-like			
CC	enzyme and encoding polynucleotides are useful for screening for			
CC	modulators which are used for treating a MAP kinase phosphatase-like			
CC	enzyme dysfunction related disease, such as asthma, a central nervous			
CC	system disorder, diabetes, obesity, chronic obstructive pulmonary			
CC	disease, cancer or a cardiovascular disease. The enzyme can be regulated			
CC	to treat allergies including asthma, allergic rhinitis, atopic			
CC	dermatitis, and anaphylaxis, central nervous system disorders such as			
CC	brain injuries, Parkinson's disease, Huntington's disease, multiple sclerosis,			
CC	stroke, Alzheimer's disease, dementia, progressive nuclear palsy, and human			
CC	disease, Creutzfeldt-Jacob dementia, and cardiovascular diseases			
CC	immunodeficiency virus (HIV) dementia, and cardiovascular diseases			
CC	including myocardial infarction, ischemic diseases of the heart, atrial			
CC	and ventricular arrhythmia, hypertensive vascular diseases and peripheral			
CC	vascular diseases. The enzyme is useful in diagnostic assays for			
CC	detecting diseases and abnormalities or susceptibility to diseases or			
CC	abnormalities related to the presence of mutations in the encoding			
CC	nucleic acid sequences. The present sequence represents the human MAP			
CC	kinase phosphatase-like enzyme polypeptide.			
XX				
XX				
SO	Sequence	779 AA;		
	Query Match	83.9%;	Score 2863;	DB 23; Length 779;
	Best Local Similarity	84.6%;	Pred. No. 8.4e-235;	
	Matches 581; Conservative	19;	Mismatches 39;	Indels 48; Gaps 11;
QY	9 SPSPGASTPV-GP-----	WDQAVQRRSRLQRRSFAVLR---	42	
DB	24 APVAGASTALWGPGRGEGTALPALTALGLLSRQDRIVQRRSLQRR---	ALRCSC	79	
QY	43 GAVTGLQDGGDNDAAEASSEPTTEKAPSEEEIH-GDQTFG-QGSQS	PQKEORQHLHL	100	
DB	80 GAVTGLQDGGDNDAAEASSEPTTEKAPSEEEIHGQTDFRCNDPRVPRSRSGSNCNL	139		
QY	101 MVQLIRPQDDIRLAQLEA-----	PRPRIRY-LIVVSTREGELSDQETVL---	LGVD	150
DB	140 MV--RRAEAGMTSDLEAQAGGTPLGIRIPCLVSVSTREGR---	RSEPRMRSSWDVD	193	
QY	151 FPDSSSPSCTGLVLPPLWSDTQVYLDGCGFSVTSGGQSRIFKPTISIQTMATLQVLHQA	210		
DB	194 FPDSSSPSCTGLVLPPLWSDTQVYLDGCGFSVTSGGQSRIFKPTISIQTMATLQVLHQA	253		

QY 211 CEALGSGLVPGSALTWASHYQERLNSEQSCINWMTAMADLESRLPPSAEBSGSSSEQE 270
DB 254 CEALGSGLVPGSALTWASHYQERLNSEQSCINWMTAMADLESRLPPSAEBSGSSSEQE 313
QY 271 MEQAIRAEIMKVLVDVSDLESVTSKEIRQALELRLGLPQYRDFIDNOMLLVAQRDRAS 330
DB 314 MERAIRAEIMKVLVDVSDLESVTSKEIRQALELRLGLPQYRDFIDNOMLLVAQRDRAS 373
QY 331 RIFPHLYLGSEWNAANLEELQNRVTHILMAREIDNEFYERFTYHNVRMLWDESAQLLP 390
DB 374 RIFPHLYLGSEWNAANLEELQNRVTHILMAREIDNEFYERFTYHNVRMLWDESAQLLP 433
QY 391 HWKETHRFIEARAQGTHTVLVHCKMGVSRSAATVLAAMKQYECGLEQALRHVQELRPIA 450
DB 434 HWKETHRFIEARAQGTHTVLVHCKMGVSRSAATVLAAMKQYECGLEQALRHVQELRPIA 493
QY 451 RPNPGFLRQLOIYQGITLTSRQSHVWEQKVGVSPEEHAPAEVSTPPPLPPEEGGEE 510
DB 494 RPNPGFLRQLOIYQGITLTSRQSHVWEQKVGVSPEEHAPAEVSTPPPLPPEEGGEE 553
QY 511 KVGMEESQAAPKEEPGRPRINIRGVMSISLLEPSLESTSETSDMEVFSSHESH 570
DB 554 KVGMEESQAAPKEEPGRPRINIRGVMSISLLEPSLESTSETSDMEVFSSHESH 613
QY 571 EEPLOPFPOLARTKGGQCVDRGFQPAKLSRQSVVTLQGSAAVAVNRTQAFQEQEGQGG 630
DB 614 EEPLOPFPOLARTKGGQCVDRGFQPAKLSRQSVVTLQGSAAVAVNRTQAFQEQEGQGG 673
QY 631 GEPCTISSTPRFRKVRQASVHDSGEEG 657
DB 674 GEPCTISSTPRFRKVRQASVHDSGEEG 700

RESULT 5
ABP51653 standard; Protein; 471 AA.
XX ABP51653;
AC ABP51653;
XX
DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 1 protein SEQ ID NO:4.
XX
KW Human; phosphatase; mitogen activated protein kinase phosphatase;
KM MAP kinase; enzyme; chromosome 11.
XX
OS Homo sapiens.
XX
PN WO200242436-A2.
XX
PD 30-MAY-2002.
XX
PF 07-NOV-2001; 2001WO-US42995.
XX
PR 20-NOV-2000; 2000US-0715177.
PR 18-JAN-2001; 2001US-0761640.
XX
PA (PEKE) PE CORP NY.
XX
PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-5/5237/61.
DR N-PSDB; ABQ73249, ABQ73252.
XX
PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies -
XX
PS Claim 1; Fig 2A; 85pp; English.
XX
CC The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting

CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC represent human MAP kinase phosphatase splice form 1 from the present
CC invention.
XX
SQ Sequence 471 AA;

Query Match 71.1%; Score 2426; DB 23; Length 471;
Best Local Similarity 100.0%; Pred. No. 7e-198;
Matches 469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALVTVSRSPPSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGGNDDAEA 60
DB 1 MALVTVSRSPPSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGGNDDAEA 60
QY 61 SSEPTEKAPSEELHGDQDFGQGSQSPQKQEQRHLMVQLRPQDDIRLAQLEAP 120
DB 61 SSEPTEKAPSEELHGDQDFGQGSQSPQKQEQRHLMVQLRPQDDIRLAQLEAP 120
QY 121 RPRRLRYLVSTREGEGLSODETVLLGVDFPDSSSPSCTLGLVPLWSDTQVYLDG 180
DB 121 RPRRLRYLVSTREGEGLSODETVLLGVDFPDSSSPSCTLGLVPLWSDTQVYLDG 180
QY 181 FSVTSGGSRIFPKPISIQTMWATLVHQAACEALGSLVPGSALTWASHYQERLNSE 240
DB 181 FSVTSGGSRIFPKPISIQTMWATLVHQAACEALGSLVPGSALTWASHYQERLNSE 240
QY 241 SCINWMTAMADLESRLPPSAEBSGSSSEQEQAIRAEIMKVLVDVSDLESVTSKEIRQAL 300
DB 241 SCINWMTAMADLESRLPPSAEBSGSSSEQEQAIRAEIMKVLVDVSDLESVTSKEIRQAL 300
QY 301 ELRLGLPQYRDFIDNOMLLVAQRDRASRI FPHLYLGSEWNAANLEELQNRVTHILN 360
DB 301 ELRLGLPQYRDFIDNOMLLVAQRDRASRI FPHLYLGSEWNAANLEELQNRVTHILN 360
QY 361 MAREIDNEFYERFTYHNVRMLWDESAQLLP HWKETHRFIEARAQGTHTVLVHCKMGVSR 420
DB 361 MAREIDNEFYERFTYHNVRMLWDESAQLLP HWKETHRFIEARAQGTHTVLVHCKMGVSR 420
QY 421 AATVLAYAMKQYECGLEQALRHVQELRPIARPNGFLRQLOIYQGITLA 469
DB 421 AATVLAYAMKQYECGLEQALRHVQELRPIARPNGFLRQLOIYQGITLA 469

RESULT 6
AAE22733 standard; Protein; 471 AA.
XX AAE22733;
AC AAE22733;
XX
DT 09-AUG-2002 (first entry)
XX
DE Murine dual-specificity phosphatase 15 (DSP-15) protein.
XX
KW Murine; dual-specificity phosphatase 15; DSP15; antiallergic; cytostatic;
KW immunosuppressive; MAP; mitogen activated protein kinase; cancer; enzyme;
KW signal transduction; cell proliferation; Duchenne muscular dystrophy;
KW cell cycle abnormality; graft-versus-host disease; autoimmune disease;
KW metabolic disease; allergy; screening.
XX
OS Mus musculus.
XX
PN WO200224740-A2.
XX
PD 28-MAR-2002.
XX
PF 19-SEP-2001; 2001WO-US29406.

XX 19-SEP-2000; 2000US-233833P.
PR 18-SEP-2001; 2001US-0955732.
XX

PA (CEPT-) CEPTYR INC.
XX

PI Luche RM, Wei B;
XX

XX WPI; 2002-394127/42.
DR N-PSDB; AAD36063.
XX

PT New dual-specificity phosphatase 15 polypeptide and polynucleotides,
PT useful for treating e.g. Duchenne muscular dystrophy, cancer,
PT graft-versus-host disease, autoimmune diseases, allergies, metabolic
PT diseases -
XX

PS Claim 50; Fig 5; 91pp; English.
XX

XX The invention relates to a new isolated dual-specificity phosphatase 15
CC (DSP-15) polypeptide which retains the ability to dephosphorylate an
CC activated MAP (mitogen activated protein) kinase. DSPs are phosphatases
CC that dephosphorylate both phosphotyrosine and phosphothreonine/serine
CC residues. DSP-15 polypeptides may be used to identify agents that
CC modulate DSP-15 activity, where such agents may inhibit or enhance signal
CC transduction via a MAP-kinase cascade, leading to cell proliferation. DSP
CC polypeptides, modulating agents, and/or polynucleotides encoding the
CC polypeptides may be used to modulate DSP-15 activity in a patient, and to
CC ameliorate disorders such as Duchenne muscular dystrophy, cancer, graft-
CC versus-host disease, autoimmune diseases, allergies, metabolic diseases,
CC abnormal cell growth, abnormal cell proliferation and cell cycle
CC abnormalities. DSP-15 alternate form polypeptides are useful in screening
CC assays for modulators of enzyme activity and/or substrate binding. The
CC present sequence is murine DSP-15 protein.
XX

SO Sequence 471 AA;

Query Match 71.1%; Score 2426; DB 23; Length 471;
Best Local Similarity 100.0%; Pred. No. 7e-198;
Matches 469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGGDNDAAEA 60
DB 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGGDNDAAEA 60

QY 61 SSEPTEKAPSEELHGDQDTDFGQSGSQSPQKOEQRQHLHWQLLRPODDIRLAOLEAP 120
DB 61 SSEPTEKAPSEELHGDQDTDFGQSGSQSPQKOEQRQHLHWQLLRPODDIRLAOLEAP 120

QY 121 RPPRLRYLLVSTREGEGLSQDETVLLGVDFPDSSSPSCITGLVPLMSDTQVYLDGDDG 180
DB 121 RPPRLRYLLVSTREGEGLSQDETVLLGVDFPDSSSPSCITGLVPLMSDTQVYLDGDDG 180

QY 181 FSVTSGGQSRIFKPISIQTMATLQVLHQAACEALGSLVPGGSALTWASHYQERLNSQ 240
DB 181 FSVTSGGQSRIFKPISIQTMATLQVLHQAACEALGSLVPGGSALTWASHYQERLNSQ 240

QY 241 SCINewTAMADLESIRPPSAEPGSSSEQMEQAIRAELWKVLDVSDLESVTSKEIRQAL 300
DB 241 SCINewTAMADLESIRPPSAEPGSSSEQMEQAIRAELWKVLDVSDLESVTSKEIRQAL 300

QY 301 ELRLGLPLQYRDFIDNOMLLLVAQRDRASRIFFHLYLGSEWNAANLELQRRNRYTHILN 360
DB 301 ELRLGLPLQYRDFIDNOMLLLVAQRDRASRIFFHLYLGSEWNAANLELQRRNRYTHILN 360

QY 361 MAREIDNFYPERFTYHNVRLWDEESAOLPHWKETHRFIEAARQGTHTLVHCKMGVRS 420
DB 361 MAREIDNFYPERFTYHNVRLWDEESAOLPHWKETHRFIEAARQGTHTLVHCKMGVRS 420

QY 421 AATVLAYAMKQYECGLEQALRHVQELRPIARPNGFLRQLOIYQGITTA 469
DB 421 AATVLAYAMKQYECGLEQALRHVQELRPIARPNGFLRQLOIYQGITTA 469

RESULT 7
ABP51655
ID ABP51655 standard; Protein; 408 AA.
XX
AC ABP51655;
XX

DT 30-SEP-2002 (first entry)
XX
DE Human MAP kinase phosphatase splice form 3 protein SEQ ID NO:6.
XX

XX Human; phosphatase; mitogen activated protein kinase phosphatase;
KW MAP kinase; enzyme; chromosome 11.
XX

OS Homo sapiens.
XX

PN WO200242436-A2.
XX

PD 30-MAY-2002.
XX

PF 07-NOV-2001; 2001WO-US42995.
XX

PR 20-NOV-2000; 2000US-0715177.
XX

PR 18-JAN-2001; 2001US-0761640.
XX

PA (PEKE) PE CORP NY.
XX

PI Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX

DR WPI; 2002-575237/61.
XX

DR N-PSDB; ABQ73255.
XX

PT Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies -
XX

PS Claim 1; Fig 2A; 85pp; English.
XX

XX The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC represent human MAP kinase phosphatase splice form 3 from the present
CC invention.
XX

SO Sequence 408 AA;

Query Match 60.1%; Score 2051.5; DB 23; Length 408;
Best Local Similarity 86.6%; Pred. No. 4.7e-166;
Matches 406; Conservative 0; Mismatches 0; Indels 63; Gaps 1;

QY 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGGDNDAAEA 60
DB 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLQDGGDNDAAEA 60

QY 61 SSEPTEKAPSEELHGDQDTDFGQSGSQSPQKOEQRQHLHWQLLRPODDIRLAOLEAP 120
DB 61 SSEPTEKAPSEELHGDQDTDFGQSGSQSPQKOEQRQHLHWQLLRPODDIRLAOLEAP 120

QY 121 RPPRLRYLLVSTREGEGLSQDETVLLGVDFPDSSSPSCITGLVPLMSDTQVYLDGDDG 180
DB 121 RPPRLRYLLVSTREGEGLSQDETVLLGVDFPDSSSPSCITGLVPLMSDTQVYLDGDDG 180

QY 181 FSVTSGGQSRIFKPISIQTMATLQVLHQAACEALGSLVPGGSALTWASHYQERLNSQ 240
DB 181 FSVTSGGQSRIFKPISIQTMATLQVLHQAACEALGSLVPGGSALTWASHYQERLNSQ 240

QY 241 SCINewTAMADLESIRPPSAEPGSSSEQMEQAIRAELWKVLDVSDLESVTSKEIRQAL 300
DB 241 SCINewTAMADLESIRPPSAEPGSSSEQMEQAIRAELWKVLDVSDLESVTSKEIRQAL 300

Db 202 -----SSSEQMEQAIRAEIWKVLDVSDLESVTSKEIRQAL 237
QY 301 ELRLGLPLQCYRDFIDNQMLLVAQRDRASRIFFPHLYLGSEWNAANLELQNRVTHILN 360
Db 238 ELRLGLPLQCYRDFIDNQMLLVAQRDRASRIFFPHLYLGSEWNAANLELQNRVTHILN 297
QY 361 MAREIDNFYPERFTYHNRLWDEBSAQLLPHWKETHRFIEARAQGTHTLVHCKMGVSR 420
Db 298 MAREIDNFYPERFTYHNRLWDEBSAQLLPHWKETHRFIEARAQGTHTLVHCKMGVSR 357
QY 421 AATVLAAYAMKQYECSLAQALRHVQELRPRIARPNGFLRQLQIYQGITLA 469
Db 358 AATVLAAYAMKQYECSLAQALRHVQELRPRIARPNGFLRQLQIYQGITLA 406

RESULT 8
ABB97419
ID ABB97419 standard; protein; 394 AA.
AC ABB97419;
DT 27-JUN-2002 (first entry)
DE Novel human protein SEQ ID NO: 687.
XX
XX Human; anti anaemic; vulnery; anti inflammatory; immunomodulator;
KW anti infertility; cerebroprotective; cytosatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR N-PSDB; ABN32605.

PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
XX
PS Claim 20; SEQ ID NO 687; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
XX
SQ Sequence 394 AA;

Query Match 55.2%; Score 1884.5; DB 23; Length 394;
Best Local Similarity 59.8%; Pred. No. 7.8e-152;
Matches 394; Conservative 0; Mismatches 0; Indels 265; Gaps 1;

QY 1 MALVTYSRSPPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLDGGDNDAAEA 60
|||||

Db 1 MALVTYSRSPPGSGASTPVGPWDQAVQRRSRLQRRQSFVLRGAVLGLDGGDNDAAEA 60
QY 61 SSEPTKAPSEELHGDQTDGQGSQSPQKQEBQRHLHLMVQLRPQDDIRLAQLAP 120
Db 61 SSEPTKAPSEELHGDQTDGQGSQSPQKQEBQRHLHLMVQLRPQDDIRLAQLAP 120
QY 121 RPPRLRYLLVSTREGEGLSQDETIVLLGVDFPDSSPSCTGLVPLPMSDTQVYLDG 180
Db 121 RPPRLRYLLVSTREGEGLSQDETIVLLGVDFPDSSPSCTGLVPLPMSDTQVYLDG 180
QY 181 FSVTSGGQSRIFKPISIQTMATLQVLHQAACEALGSLVPGSALTWASHYQERLNS 240
Db 181 FSVTSGGQSRIFKPISIQTMATLQVLHQAACEALGSLVPGSALTWASHYQERLNS 240
QY 241 SCINWTAMADLESIRPPSAPSGSSEQEOMEQAIRAELMKVLDVSDLESVTSKEIR 300
Db 241 SCINWTAMADLESIRPPSAPSGSSEQEOMEQAIRAELMKVLDVSDLESVTSKEIR 282
QY 301 ELRLGLPLQCYRDFIDNQMLLVAQRDRASRIFFPHLYLGSEWNAANLELQNRVTHILN 360
Db 283 -----
QY 361 MAREIDNFYPERFTYHNRLWDEBSAQLLPHWKETHRFIEARAQGTHTLVHCKMGVSR 420
Db 283 -----
QY 421 AATVLAAYAMKQYECSLAQALRHVQELRPRIARPNGFLRQLQIYQGITLASRQSHWEQKV 480
Db 283 -----
QY 481 GGVSPPEHPAPEVSTPFPPLPEPEGGEKVGMEESQAAPKEPGRPRINLRGVMRS 540
Db 283 -----
QY 541 ISLLEPSLESTSETSDMPEVFSHSHSHEEPLOPPOLARTKGGQVDRGPQALKSR 600
Db 283 -----LELESTSETSDMPEVFSHSHSHEEPLOPPOLARTKGGQVDRGPQALKSR 335
QY 601 QSVVTLQGSAAVVAANTQAFQEQEQGGGQGGEPCTISSTPRFRKVVRRQASVHDSGEEGEA 659
Db 336 QSVVTLQGSAAVVAANTQAFQEQEQGGGQGGEPCTISSTPRFRKVVRRQASVHDSGEEGEA 394

RESULT 9
AAB73226
ID AAB73226 standard; protein; 341 AA.
XX
AC AAB73226;
XX
DT 11-MAY-2001 (first entry)
XX
DE Human phosphatase NP_060746_h.
XX
DE
XX
KW Phosphatase; gene therapy; enzyme; cancer; pathophysiological hypoxia;
KW cardiac dysfunction; vascular disorder; myopathy; ectodermal dysplasia;
KW congenital muscle disorder; Papillon-Lefevre syndrome; Cowden disease;
KW Moebius syndrome; Bjornstad syndrome; Bannayan Zonana syndrome;
KW schizophrenia; hamartoma.
XX
OS Homo sapiens.
XX
PN WO200112819-A2.
XX
PD 22-FEB-2001.
XX
PF 11-AUG-2000; 2000WO-US22158.
XX
PR 13-AUG-1999; 99US-0149005.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Hill R, Flanagan P, Lioubin M;

DR WPI; 2001-211226/21.
DR N-PSDB; AAF63578.
XX
XX New protein phosphatase polypeptide for diagnosing and treating
PT phosphatase related disorders such as cancer, schizophrenia, cardiac
PT dysfunction and/or vascular disorders
XX
XX Claim 6; Fig 5; 138pp; English.
XX
CC The present invention relates to phosphatase proteins and coding
CC sequences. The present sequence is one such phosphatase. Phosphatases are
CC enzymes that catalyze the dephosphorylation of proteins modified by
CC phosphorylation of serine, threonine or tyrosine residues. The
CC phosphatases are useful for treating a variety of diseases: for example
CC cancer e.g. breast, uterine, prostate, head, neck, lung cancers,
CC synovial sarcoma, renal cell carcinoma, non-small cell lung cancer,
CC hepatocellular carcinoma, pancreatic endocrine tumours, stomach cancer,
CC glioblastoma, colorectal cancer and thyroid cancer, pathophysiological
CC hypoxia, cardiac dysfunction and/or vascular disorders, myopathies,
CC congenital muscle disorders, Papillon-Lefevre syndrome, Cowden disease,
CC ectodermal dysplasia, Moebius syndrome, Bjornstad syndrome, Bannayan
CC Zonana syndrome, schizophrenia and hamartomas.
XX
SQ Sequence 341 AA;
Query Match 52.1%; Score 1779; DB 22; Length 341;
Best Local Similarity 100.0%; Pred. No. 6.1e-143;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 319 MLLVAQRDRASRIFPHLYLGSEWMAANLELQNRVTHILNMAEIDNFBERTYHNV 378
DB 1 MLLVAQRDRASRIFPHLYLGSEWMAANLELQNRVTHILNMAEIDNFBERTYHNV 60
QY 379 RLWDEESAQLPFWKETHRFIEAARQGTHTLVHCKMGVSRSAATVLAAMKQYECSE 438
DB 61 RLWDEESAQLPFWKETHRFIEAARQGTHTLVHCKMGVSRSAATVLAAMKQYECSE 120
QY 439 ALRHVQELRPIARPNGFLRQLQIYQILTASRQSHWEEKVGVSPPEHPAPEVSTPP 498
DB 121 ALRHVQELRPIARPNGFLRQLQIYQILTASRQSHWEEKVGVSPPEHPAPEVSTPP 180
QY 499 PLPPEEGGEEKVGMESQAAPKEBGP RPRLNLRGVMSISLLEPSTLESTSETSD 558
DB 181 PLPPEEGGEEKVGMESQAAPKEBGP RPRLNLRGVMSISLLEPSTLESTSETSD 240
QY 559 MPEVSSHESHHEPLQPPQARTKGQVDRGPQALKSRQSVTLQGSAAVAVNRTOA 618
DB 241 MPEVSSHESHHEPLQPPQARTKGQVDRGPQALKSRQSVTLQGSAAVAVNRTOA 300
QY 619 FOEQEGQGQGQGEPCISSTPRFRKVVROASVHDSGEEGEA 659
DB 301 FOEQEGQGQGQGEPCISSTPRFRKVVROASVHDSGEEGEA 341
RESULT 10
ID ABP51656 standard; Protein; 312 AA.
XX ABP51656;
XX AC
XX 30-SEP-2002 (first entry)
XX DT
XX Human MAP kinase phosphatase related protein sequence SEQ ID NO:8.
XX DE
XX Human; phosphatase; mitogen activated protein kinase phosphatase;
XX KW MAP kinase; enzyme.
XX OS Homo sapiens.
XX XX
XX WO200242436-A2.
XX PN
XX 30-MAY-2002.
XX PD
XX

PF 07-NOV-2001; 2001WO-US42995.
XX
XX 20-NOV-2000; 2000US-0715177.
PR 18-JAN-2001; 2001US-0761640.
XX
XX (PEKE) PE CORP NY.
XX
XX Wei M, Ketchum KA, Di Francesco V, Beasley EM;
XX
XX WPI; 2002-575237/61.
XX
XX Novel isolated human phosphatase peptide useful for treating disorder
PT characterized by absence of, inappropriate or unwanted expression of
PT the phosphatase protein, and as immunogens to raise antibodies
XX
XX Disclosure; Fig 2C; 85pp; English.
XX
XX The present invention describes an isolated human phosphatase peptide
CC (I). (I) can be used for identifying a modulator of (I) by contacting
CC (I) with an agent and determining if the agent has modulated the
CC function or activity of (I). (I) is useful for identifying an agent that
CC binds to (I), by contacting (I) with an agent and assaying the contacted
CC mixture to determine whether a complex is formed with the agent bound
CC (I). The human phosphatases from the present invention are mitogen
CC activated protein (MAP) kinase phosphatases. These human MAP kinase
CC phosphatases are located on chromosome 11. (I) and the polynucleotide
CC sequences encoding (I) can be used in gene therapy. The present sequence
CC represents a human protein which is given in comparison with a human MAP
CC kinase phosphatase from the present invention.
XX
SQ Sequence 312 AA;
Query Match 43.1%; Score 1470.5; DB 23; Length 312;
Best Local Similarity 92.9%; Pred. No. 1.1e-116;
Matches 290; Conservative 6; Mismatches 15; Indels 1; Gaps 1;
QY 1 MALVTVSRSPPSGASTPVGPWDQAVQRRSLQRRQSFVAVLRGAVLGLQDGGNDDAEA 60
DB 1 MALVTVSRSPPSGASTPVGPWDQAVQRRSLQRRQSFVAVLRGAVLGLQDGGNDDAEA 60
QY 61 SSEPTEKAPSEBELHGDQDTDFGQGSQSPQKEQRQHLHNVQLRPDDIRLAAQLAP 120
DB 61 SSEPTEKAPSEBELHGDQDTDFGQGSQSPQKEQRQHLHNVQLRPDDIRLAAQLAP 120
QY 121 RPPRLRYLVVSTREGEGLSQDETIVLGVDFPDSSSPCTGLVLPWSDTQVYLDG 180
DB 121 RPPRLRYLVVSTREGEGLSQDETIVLGVDFPDSSSPCTGLVLPWSDTQVYLDG 180
QY 181 FSVTSGGQSRIFKPIISIQTMWATLQVLAQCEALGSGIVPGSALTWASHYQERINSEQ 240
DB 181 FSVTSGGQSRIFKPIISIQTMWATLQVLAQCEALGSGIVPGSALTWASHYQERINSEQ 240
QY 241 SCINWETAMADLESILRPPSAEPGSSSQEQMEQAIRAELMKVLDV-SLSEVTSKEIRQA 299
DB 241 SCINWETAMADLESILRPPSAEPGSSSQEQMEQAIRAELMKVLDV-SLSEVTSKEIRQA 300
QY 300 LELRLGLPLQY 311
DB 301 HESSHEEPLQPF 312
RESULT 11
ID AAE04833 standard; Protein; 1049 AA.
XX AAE04833;
XX AC
XX 10-SEP-2001 (first entry)
XX DT
XX Human SGP006 phosphatase polypeptide.
XX DE
XX Human; SGP006 phosphatase polypeptide; phosphatase-related disease;
XX KW immune-related disorder; ocular disease; organ transplant rejection;
XX

KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiant;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cyostatic;
KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnerary; tranquilliser; antiasthmatic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome 12q21.3-q22.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..173
FT /label= Catalytic_domain
FT Domain 308..446
FT /label= Phosphatase_domain
XX
PN WO200146394-A2.
XX
PD 28-JUN-2001.
XX
PF 21-DEC-2000; 2000WO-US34736.
XX
PR 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JAN-2000; 2000US-0179301.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ,
PI Flanagan P;
XX
DR WPI; 2001-418058/44.
DR N-PSDB; AAD09491.
XX
XX
PT Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders -
XX
PS Claim 7; Fig 2; 186pp; English.
XX
CC The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC haematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infections caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognitive
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders, dyskinesias and organ transplant rejection. The present
CC amino acid sequence is human SGP006 phosphatase polypeptide. This
CC sequence is classified as dual specificity phosphatase (DSP) and MAP
CC kinase phosphatase (MKP). SGP006 gene maps to chromosomal position
CC 12q21.3-q22.
XX
SQ Sequence 1049 AA;
Query Match 34.3%; Score 1172; DB 22; Length 1049;
Best Local Similarity 42.3%; Pred. No. 2e-90;
Matches 269; Conservative 95; Mismatches 170; Indels 102; Gaps 18;
QY 1 MALVTVSRSPPSGASTPVGPMW-----QAVQRRSLQRQSFVLRGAVLGLDGGDND 56
DB 1 MALVTLQRRSPPTSAASSASNSLEAGSEEDRKLNLSSESFVMVKGALFLQGG----- 55

QY 57 AABASSEPTKAPSEELHGDQTPDQGS-QSPQKQ-EQRQHLIMVQLRPDDIRLA 114
DB 56 -----SSPQ-----GQSLQHPKHAGDLPQHLQVMINLRCEDRIKLA 94
QY 115 AQLAPRPRIARLLV--STREGGLSDQETVLLGVDPEDSSPSCTLGLVPLMSDTQ 172
DB 95 VRLESAMADRVYVWVYSSGRQ---DTEENILLGVDPSSKESKSCCTIGVLRIMSDTK 150
QY 173 VYLDGDFGSVYSGQSRIFKPIQTMWATLOVLHQAALGSLVPGSALTWASHY 232
DB 151 IHLDDGGSVSTAGRMHIFKPVSVQAMWSALQVLHKAQVARRHNYFPGVALIWTATY 210
QY 233 QERLNSQSCINEMWTAMADLESRP--PSAEPGSSSEQOMEQAIRAELMKVLDVSDLES 290
DB 211 ESCISSEQSCINEMWAMQDLESTRDPSPALFVDKPTGEERTERLIKAKLRSIMSQDLEN 270
QY 291 VTSKEIRQALRLGLPLQYRDPIDNQMLLLVAQRDRASRIFPHLYLGEWNAANLEEL 350
DB 271 VTSKEIRNELKQMNCKLKEFDNEMLLILGQMDKPSLIFDLYLGSEWNASNLEEL 330
QY 351 QNRVTHILNMAREIDNFFPERFYANVRLMDEESAQLPHWKETHRFIEAPARQGTHTVL 410
DB 331 QSGVDYILNVTREIDNFFPGLFAYHNIRVYDEETDLDLAHWNAYHFINKAKRNSKCL 390
QY 411 VHCKMGVSRSAATVLYAMKQYECGLQALRHVQELRIARPNGFLROLQYOGILITAS 470
DB 391 VHCKMGVSRSAATVLYAMKEFGWPLEKAYNVKQKRSITRPNAGFMROLSEYEGILDAS 450
QY 471 RQSH--VWEQYVG--VSPEEHPA-----PEVSTPF-----PPLP-- 501
DB 451 KQRHNLKWRQQTDSLSLQCPVDDPAGPDPLPETPDGTPESQLPFLDDAQPGLPPLPCC 510
QY 502 -----PEPEGGEKRVGMEE-----SQAPKEE---PGRPRINLRGVMS 540
DB 511 FRRLSDPLLPSPEDTG--SLVHLEDPERALLFEAPPAEVHRPARQPO-----QG 560
QY 541 ISLLEPSL-ELESTSETSDMPEVSSHESHEEPL 574
DB 561 SGLCEKDVKKLFFGSPKRGSGLLQVEETEREEL 596
RESULT 12
AAE04835
ID AAE04835 standard; Protein; 498 AA.
XX
AC AAE04835;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human SGP001 phosphatase polypeptide.
XX
KW Human; SGP001 phosphatase polypeptide; phosphatase-related disease;
KW immune-related disorder; ocular disease; organ transplant rejection;
KW infection; diabetes; pain; sexual dysfunction; Alzheimer's disease;
KW metabolic disorder; haematopoietic cancer; mood disorder; cardiant;
KW Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW cardiovascular disease; brain; neuronal-associated disease; dyskinesia;
KW attention disorder; cognition disorder; psychotic disorder; cyostatic;
KW neurological disorder; virucide; nootropic; cerebroprotective; therapy;
KW neuroprotective; antibacterial; vulnerary; tranquilliser; antiasthmatic;
KW hypotensive; immunosuppressive; antipsoriatic; analgesic; hypertensive;
KW antifungal; dual specificity phosphatase; DSP; MAP kinase phosphatase;
KW MKP; migraine; chromosome Xp11.1-11.3.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..173
FT /label= Catalytic_domain
FT Domain 307..441
FT /label= Phosphatase_domain
XX
PN WO200146394-A2.

```

XX 28-JUN-2001.
PD
XX
PF 21-DEC-2000; 2000WO-US34736.
XX
XX 21-DEC-1999; 99US-0173255.
PR 28-DEC-1999; 99US-0175766.
PR 25-JAN-2000; 2000US-0178078.
PR 31-JAN-2000; 2000US-0179301.
XX
PA (SUGEN-) SUGEN INC.
XX
PI Plowman GD, Martinez R, Whyte D, Manning G, Sudarsanam S, Hill RJ;
PI Flanagan P;
XX
DR WPI; 2001-418058/44.
DR N-PSDB; AAD09493.
XX
XX Novel phosphatase polypeptide useful for treating cancers,
PT immune-related diseases and disorders, cardiovascular disease, brain or
PT neuronal-associated diseases and metabolic disorders
XX
PS Claim 7; Fig 2; 186pp; English.
XX
XX The present invention relates to phosphatase polypeptides, nucleotide
CC sequences encoding them, as well as various products and methods useful
CC for the diagnosis and treatment of various phosphatase-related diseases
CC and conditions. Substance that modulates the activity of phosphatase
CC polypeptide is used to treat immune-related diseases and disorders,
CC cardiovascular disease, brain or neuronal-associated diseases and
CC metabolic disorders, including cancers of tissues, cancers of
CC hematopoietic origin, diseases of central and peripheral nervous
CC system, Alzheimer's disease, Parkinson's disease, multiple sclerosis,
CC amyotrophic lateral sclerosis, viral infections, infectious caused by
CC prions, bacteria and fungi, ocular diseases, diabetes, migraines, pain,
CC sexual dysfunction, mood disorders, attention disorders, cognitive
CC disorders, hypotension, hypertension, psychotic disorders, neurological
CC disorders, dyskinesias and organ transplant rejection. The present
CC amino acid sequence is human GSP001 phosphatase polypeptide. This
CC sequence is classified as dual specificity phosphatase (DSP) and MAP
CC kinase phosphatase (MKP). GSP001 gene maps to chromosomal position
CC Xp11.1-11.3.
XX
SQ Sequence 498 AA;

```

Query Match 31.9%; Score 1089.5; DB 22; Length 498;
 Best Local Similarity 42.1%; Pred. No. 6.7e-84;
 Matches 227; Conservative 102; Mismatches 149; Indels 61; Gaps 9;

```

QY 1 MALVTVSRSPGSGASTPVGPMQAVQRRRLQRR--QSPAVLRGAVLGLQDGDNDDA 57
DB 1 MALVTVQSRPTPTSTSSPCASEADSGEEERSQPRISSEFLVKGALLFLPR----- 53
QY 58 AEASSEPTKAPSEELHGDQDTFGQSGSQSPQKEBQ-----RQHILMVQLRPQDD 110
DB 54 -----GNGSSTPRISHRNKHAQDLQHLQAMFILRPEDN 89
QY 111 IRLAQLAEPRLPRRLXLLVSTREGELSDQDETLLGVDFPDSSSPSCITGLVPLMSD 170
DB 90 IRLAVRLSTYQNRTRYMVVSTNGRQ--DTEESIVLGMDFSSNDSTCTMGLVPLMSD 147
QY 171 TQVYLDGDDGFSVTSQSGSRIKPISTQTMWATLQVLHQAQCEALGSGLVPGSGSLTWS 230
DB 148 TLIHLDGDDGFSVSTINRVIHFQPSVQAMWSALQSLHKAQCEVARAHNYPPGSLFTWS 207
QY 231 HYQERLNSQSCLENWTAMADLESRLP--PSAEPGSSSQEQMEQAIRAELKVLVDVSD 288
DB 208 YYESHINSQSSVNEWMAMQDVQSHRPDPSALFTDIPTERERTERLIKTKLREIMWQXL 267
QY 289 ESVTSKEIRQALEIRGLPLQCYRDFIDNOMLLVAQRDRASRIFFHLVYLGSEWNAANLE 348
DB 268 ENITSKEIRTELEMQVNCNLRFPKEFIDNEMIVILGQMDSPQIFEHVFLGSEWNAANLE 327

```

```

QY 349 ELQNRNVTHTILNMAREIDNFFPERFTYHNVRLWDEESAQLBHWKETRFFIEAARAQGT 408
DB 328 DLQNRGVRYTILNVTREIDNFFPGVFVEYHNI RVYDEEATDLAYWNDTYKFI SKAKHGSK 387
QY 409 VLVHCKMGVSRSAATVLAAYAMKQYECGLEQALRHVQELRPAPNPNGFLRQLQYOGILT 468
DB 388 CLVHCKMGVSRSAATVLAAYAMKEY----DRAYDYKERRRTVAKPNPSFMRQLBEYOGILL 443
QY 469 ASRQSHVWEQXVGVSPEEHAPAEVSTPRP-----PLPPEPGGGEKVGVMESQAAP 522
DB 444 ASFLGLIH---GG---RDKPWEKSTEFESVDLVISPGSCCNPEKLIHISHPYLTP 495

```

RESULT 13

AAE06775

ID AAE06775 standard; Protein; 509 AA.

AC AAE06775;

DT 16-OCT-2001 (first entry)

DE Human dual-specificity phosphatase (DSP)-13 protein.

Human; dual-specificity phosphatase; DSP-13; dual phosphorylation motif; mitogen-activated protein kinase; MAP-kinase; graft-versus-host disease; GVHD; Duchenne muscular dystrophy; therapy; cancer; autoimmune disease; allergy; metabolic disease; cell growth; cell proliferation; cytostatic; cell cycle abnormality; cell differentiation; anti-allergic; muscular; immunosuppressive.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Domain 395..415

PN WO200157221-A2.

PD 09-AUG-2001.

PF 01-FEB-2001; 2001WO-US03429.

PR 02-FEB-2000; 2000US-0179886.

PA (CEPT-) CEPTYR INC.

PI Luche RM, Wei B;

DR WPI; 2001-488887/53.

DR N-PSDB; AAD12966.

New isolated dual-specificity phosphatase polypeptide for treating cancer, graft-versus-host disease, autoimmune diseases, allergies, metabolic diseases, abnormal cell growth and abnormal cell proliferation

PS Claim 50; Fig 4; 81pp; English.

The present sequence is human dual-specificity phosphatase (DSP)-13 protein. Inactivation of mitogen-activated protein kinase (MAP-kinase) is mediated by dephosphorylation at a dual phosphorylation motif by DSP which is referred to as MAP-kinase phosphatase. An agent that modulates DSP is useful for treating a disorder selected from Duchenne muscular dystrophy, cancer, graft-versus-host disease (GVHD), autoimmune diseases, allergies, metabolic diseases, abnormal cell growth, abnormal cell proliferation and cell cycle abnormalities. DSP is useful for identifying antibodies and other agents that inhibit DSP-12 and/or DSP-13 activity. DSP and the agents identified are useful for modulating cell proliferation, differentiation and survival. DSP is useful in screening assays for modulators of enzyme activity and substrate binding and for dephosphorylating a substrate of DSP-12 or DSP-13.

SQ	Sequence	509 AA;	
	Query Match	31.0%; Score 1058; DB 22; length 509;	
	Best Local Similarity	42.5%; Pred. No. 3.4e-81;	
	Matches	213; Conservative 100; Mismatches 134; Indels 54; Gaps 7;	
QY	36	QSFAYLRGAVLGLDGDNDDAEASSEPTEKAPSEELHGDQDTDFGQGSQSPQKQEQ-	94
Db	46	ESFLTVKGAALFLPR-----GNGSSTPRISHRRN	74
QY	95	-----RQHLHLMVQLRPQDRIRLAQLAPRPPRLRYLLVSTREGESLSQDETVLIG	148
Db	75	KHAGDLQOHLQAMFILLRPEDNIRLAVRLESTYQNRTRYMVVSTNGRQ--DTEESTVLG	132
QY	149	VDPPDSSPSCTGLVLPPLMSDTQVYLDGDFSVTSGGQSRIFKPISTQTMATLQVLH	208
Db	133	MDPSSNDSTCTMGLVLPPLMSDTLHLDGDFSVSTDNRVHIFKPVSVQAMWSALQSLH	192
QY	209	QACEALGSGLVFGGSALTWASHYQERLNSQSCLENTAMADLESIRP--PSAEPGGSS	266
Db	193	KACEVARAHNYYPGSLFLTWVSYESHINSQSSVNEWMAQDVQSHRPDSPALFTDIP	252
QY	267	EQEQMEQAIRAELMKVLDVSDLESVTSKEIRQALERLGLPLQQRDFIDNQMLLVAQR	326
Db	253	ERERTERLIKTKLREIMQKDLNITSKEIRTELEMQVNCNIREKFEFIDNEMIVILQGM	312
QY	327	DRASRIFPHLYLGSEWNAANLEIQRNVTILNMAREIDNFPERTYHNRLWDEESA	386
Db	313	DSPTQIFEHVFLGSEWNASNLEDLQNRGVRYILNVTREIDNFPFVFHEYHNIRVYDEAT	372
QY	387	QLLPHWKETHREIEARAQGTHTVLVCKMGVSRSAATVLAAMKQYECSLQALRHVOEL	446
Db	373	DLAYWMDTYKFTISKAKHGSKCLVCKMGVSRSAATVLAAMKEYGWNLDRAVDYKER	432
QY	447	RPIARPNGFLRQIQYOGILTASRQSHWQKVGVSPEEHPAPEVSTPFP-----PLP	501
Db	433	RTVTKPNPSFMRQLEBYOGILLASFLGLIH---GG---RDKPWEKSTEFESVDLVISIP	485
QY	502	PEPEGGGEKVVGMESQAAP	522
Db	486	GSPSCCNPEKLLHISHPYLTP	506
RESULT 14			
AAE07044			
ID	AAE07044 standard; Protein; 509 AA.		
XX	AC	AAE07044;	
XX	DT	16-OCT-2001 (first entry)	
XX	Human dual-specificity phosphatase (DSP)-13 mutant protein, D368A.		
XX	Human; dual-specificity phosphatase; DSP-13; dual phosphorylation motif;		
KM	mitogen-activated protein kinase; MAP-kinase; graft-versus-host disease;		
KM	GVHD; Duchenne muscular dystrophy; therapy; cancer; autoimmune disease;		
KM	allergy; metabolic disease; cell growth; cell proliferation; cytostatic;		
KM	cell cycle abnormality; cell differentiation; antiallergic; muscular;		
KM	immunosuppressive; mutant; mutein; variant.		
XX	OS	Homo sapiens.	
OS	OS	Synthetic.	
XX	Key	Location/Qualifiers	
FT	Misc-difference	368	
FT	/note= "Wild type Asp substituted with Ala"		
XX	WO200157221-A2.		
PN			
XX			
PD	09-AUG-2001.		
XX			
PF	01-FEB-2001; 2001WO-US03429.		
XX			

PR	02-FEB-2000; 2000US-0179886.	
XX	PA	(CEPT-) CEPTYR INC.
XX	PI	Luche RM, Wei B;
XX	XX	WPI; 2001-488887/53.
DR	New isolated dual-specificity phosphatase polypeptide for treating	
PT	cancer, graft-versus-host disease, autoimmune diseases, allergies,	
PT	metabolic diseases, abnormal cell growth and abnormal cell	
PT	proliferation	
XX	Disclosure; Page -; 81pp; English.	
PS	The present sequence is human dual-specificity phosphatase (DSP)-13	
CC	mutant protein, D368A. Inactivation of mitogen-activated protein kinase	
CC	(MAP-kinase) is mediated by dephosphorylation at a dual phosphorylation	
CC	motif by DSP which is referred to as MAP-kinase phosphatase. An agent	
CC	that modulates DSP is useful for treating a disorder selected from	
CC	Duchenne muscular dystrophy, cancer, graft-versus-host disease (GVHD),	
CC	autoimmune diseases, allergies, metabolic diseases, abnormal cell growth,	
CC	abnormal cell proliferation and cell cycle abnormalities. DSP is useful	
CC	for identifying antibodies and other agents that inhibit DSP-12 and/or	
CC	DSP-13 activity. DSP and the agents identified are useful for	
CC	modulating cell proliferation, differentiation and survival. DSP is	
CC	useful in screening assays for modulators of enzyme activity and	
CC	substrate binding and for dephosphorylating a substrate of DSP-12 or	
CC	DSP-13.	
CC	Note: The present sequence is not shown in the specification, but is	
CC	derived from the DSP-13 sequence (AAE06775) given as SEQ ID NO: 6 in	
CC	figure 4.	
XX	Sequence	
SO	509 AA;	
Query Match		
Best Local Similarity 30.8%; Score 1050; DB 22; length 509;		
Matches 212; Conservative 100; Mismatches 135; Indels 54; Gaps 7		
QY	36	QSFAYLRGAVLGLDGDNDDAEASSEPTEKAPSEELHGDQDTDFGQGSQSPQKQEQ-
Db	46	ESFLTVKGAALFLPR-----GNGSSTPRISHRRN
QY	95	-----RQHLHLMVQLRPQDDIRLAQLAPRPPRLRYLLVSTREGESLSQDETVLLG
Db	75	KHAGDLQOHLQAMFILLRPEDNIRLAVRLESTYQNRTRYMVVSTNGRQ--DTEESIVLG
QY	149	VDPPDSSPSCTGLVLPPLMSDTQVYLDGDFSVTSGGQSRIFKPISTQTMATLQVLH
Db	133	MDPSSNDSTCTMGLVLPPLMSDTLHLDGDFGSVSTDNRVHIFKPVSVQAMWSALQSLH
QY	209	QACEALGSGLVFGGSALTWASHYQERLNSQSCLENTAMADLESIRP--PSAEPGGSS
Db	193	KACEVARAHNYYPGSLFLTWVSYESHINSQSSVNEWMAQDVQSHRPDSPALFTDIP
QY	267	EQEQMEQAIRAELMKVLDVSDLESVTSKEIRQAELRLGLPLQQRDFIDNQMLLVAQR
Db	253	ERERTERLIKTKLREIMQKDLNITSKEIRTELEMQVNCNIREKFEFIDNEMIVILQGM
QY	327	DRASRIFPHLYLGSEWNAANLEIQRNVTILNMAREIDNFPERTYHNRLWDEESA
Db	313	DSPTQIFEHVFLGSEWNASNLEDLQNRGVRYILNVTREIDNFPFVFHEYHNIRVYAEAT
QY	387	QLLPHWKETHREIEARAQGTHTVLVCKMGVSRSAATVLAAMKQYECSLQALRHVOEL
Db	373	DLAYWMDTYKFTISKAKHGSKCLVCKMGVSRSAATVLAAMKEYGWNLDRAVDYKER
QY	447	RPIARPNGFLRQIQYOGILTASRQSHWQKVGVSPEEHPAPEVSTPFP-----PLP
Db	433	RTVTKPNPSFMRQLEBYOGILLASFLGLIH---GG---RDKPWEKSTEFESVDLVISIP
QY	502	PEPEGGGEKVVGMESQAAP

Db 486 GSPSCNPEKLIHSHPYLTP 506

RESULT 15

AAE07045

ID AAE07045 standard; Protein; 509 AA.

XX AAE07045;

AC AAE07045;

XX 16-OCT-2001 (first entry)

XX Human dual-specificity phosphatase (DSP)-13 mutant protein, C399S.

DE Human; dual-specificity phosphatase; DSP-13; dual phosphorylation motif; mitogen-activated protein kinase; MAP-kinase; graft-versus-host disease; GVHD; Duchenne muscular dystrophy; therapy; cancer; autoimmune disease; allergy; metabolic disease; cell growth; cell proliferation; cytosolic; cell cycle abnormality; cell differentiation; antiallergic; muscular; immunosuppressive; mutant; mutein; variant.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 399 /note= "Wild type Cys substituted with Ser"

FT

XX WO200157221-A2.

PN

XX 09-AUG-2001.

PD

XX 01-FEB-2001; 2001WO-US03429.

PF

XX 02-FEB-2000; 2000US-0179886.

PR

XX (CEPT-) CEPTYR INC.

PA

XX Luche RM, Wei B;

PI

XX WPI; 2001-488887/53.

DR

XX New isolated dual-specificity phosphatase polypeptide for treating cancer, graft-versus-host disease, autoimmune diseases, allergies, metabolic diseases, abnormal cell growth and abnormal cell proliferation

PT

XX Disclosure; Page -: 81pp; English.

PS

XX The present sequence is human dual-specificity phosphatase (DSP)-13 mutant protein, C399S. Inactivation of mitogen-activated protein kinase (MAP-kinase) is mediated by dephosphorylation at a dual phosphorylation motif by DSP which is referred to as MAP-kinase phosphatase. An agent that modulates DSP is useful for treating a disorder selected from Duchenne muscular dystrophy, cancer, graft-versus-host disease (GVHD), autoimmune diseases, allergies, metabolic diseases, abnormal cell growth, abnormal cell proliferation and cell cycle abnormalities. DSP is useful for identifying antibodies and other agents that inhibit DSP-12 and/or DSP-13 activity. DSP and the agents identified are useful for modulating cell proliferation, differentiation and survival. DSP is useful in screening assays for modulators of enzyme activity and substrate binding and for dephosphorylating a substrate of DSP-12 or DSP-13.

CC Note: The present sequence is not shown in the specification, but is derived from the DSP-13 sequence (AAE06775) given as SEQ ID NO: 6 in figure 4.

CC

XX

SQ Sequence 509 AA;

Query Match 30.7%; Score 1048; DB 22; Length 509;

Best Local Similarity 42.3%; Pred. No. 2.4e-80;

Matches 212; Conservative 100; Mismatches 135; Indels 54; Gaps 7;

QY 36 QSAFVLRGAVLGLQDGDNDDAEASSEPTEKAPSEELHGDQTFDGGGSGSPQKQEEQ- 94

Db 46 ESFLTVKGAALFLPR-----GNGSSTPRISHRRN 74

QY 95 -----RQHLHLMVQLRPODDIRLAQLEAPRPPRLRYLLVSTREGGLSODETVLLG 148

Db 75 KHAGDLQQLQAMFILRPEDNIRLAVRLESTYQNRTRYMVVSTNGRQ--DTEESTVLG 132

QY 149 VDFPDSSSPSCTLGLVPLMSDQVYLDGDFSVTSGGQSRIFKPISTQTMWATLQVLH 208

Db 133 MDFSNDSSSTCTMGLVPLMSDQVYLDGDFSVTSNDNRVHIFKPVSVQAMWSALQSLH 192

QY 209 QACEAALGSGLVPGGSAITWASHYQERLNSRQSCLENWTAMADLESRLP--PSAEPGGSS 266

Db 193 KACEVARRAHNYPPGSLFTWVSYESHINSQSSVNEWMAMQDVQSHRPPDSPALFTDIP 252

QY 267 EQQMEQAIRAELWKVLDVSDLESVTSKEIRQALERLGLPLQYRDFIDNQLLVAAQR 326

Db 253 ERERTERLIKTLREIMMOKDLENITSKEIRTELEMQVNCNLRERKEFIDNEMIVILGQM 312

QY 327 DRASRIFFHLYLGEWNAANLEELQNRNVTHTLMMAREIDNEFYPERFTYHNRVLDDESA 386

Db 313 DSPTQIFEHVFLGSEWNASNLEDLQNRGVRYILNVTREIDNFPQGVFEYHNRVYDEEAT 372

QY 387 QLEPWKETHRFIEAARAQGTHTLVHCKMGVSRSAATVLAAMQVECSLEQALRHVDEL 446

Db 373 DLAYWMDTYKFKISKAKHGSKCLVHSMGVSRSASTVIAYAMKEYGWNLDRAYDYKER 432

QY 447 RPIARPNGFLRLQIYQILITASRQSHVMEQKVGVSPEEHAPAEVSTPP-----PLP 501

Db 433 RTVTKPNPSFMRQLBEYQILLASFLGLIH---GG---RDKPGEKSTEFESVDLVSTIP 485

QY 502 PEPEGGEEKVVGMEBSQAAP 522

Db 486 GSPSCNPEKLIHSHPYLTP 506

Search completed: January 15, 2004, 06:48:27

Job time : 76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 06:49:26 ; Search time 39 Seconds

(without alignments)
3455.208 Million cell updates/sec

Title: US-09-955-732-2
Perfect score: 3412
Sequence: 1 MALVTVSRRSPGSGASTPVG.....RFRKVVRQASVHDSGEEGEA 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3412	100.0	659	US-09-955-732-2	Sequence 2, Appli
2	3402	99.7	659	US-09-761-640-5	Sequence 5, Appli
3	2863	83.9	779	US-10-363-676-11	Sequence 11, Appli
4	2623	76.9	513	US-10-108-260A-2517	Sequence 2517, Ap
5	2426	71.1	471	US-09-761-640-4	Sequence 4, Appli
6	2426	71.1	471	US-09-955-732-21	Sequence 21, Appli
7	2051.5	60.1	408	US-09-761-640-6	Sequence 6, Appli
8	1470.5	43.1	312	US-09-761-640-8	Sequence 8, Appli
9	1136	33.3	703	US-10-108-260A-3142	Sequence 3142, Ap
10	1058	31.0	509	US-09-775-925-6	Sequence 6, Appli
11	1058	31.0	509	US-09-955-732-14	Sequence 14, Appli
12	1005	29.5	484	US-10-181-590-9	Sequence 9, Appli
13	971	28.5	552	US-09-775-925-2	Sequence 2, Appli
14	971	28.5	552	US-09-955-732-15	Sequence 15, Appli
15	918.5	26.9	524	US-09-761-640-9	Sequence 9, Appli

16	914.5	26.8	737	10	US-09-955-732-13	Sequence 13, Appli
17	911	26.7	986	12	US-10-381-333-7	Sequence 7, Appli
18	581.5	17.0	140	10	US-09-955-732-12	Sequence 12, Appli
19	581.5	17.0	141	12	US-10-363-676-2	Sequence 2, Appli
20	578	16.9	111	10	US-09-761-640-10	Sequence 10, Appli
21	574	16.8	241	9	US-09-775-925-8	Sequence 8, Appli
22	539	15.8	244	12	US-10-240-145-81	Sequence 81, Appli
23	539	15.8	244	15	US-10-233-131-29	Sequence 29, Appli
24	521	15.3	170	9	US-09-775-925-31	Sequence 31, Appli
25	512	15.0	170	9	US-09-775-925-32	Sequence 32, Appli
26	497	14.6	458	15	US-10-103-313-362	Sequence 362, App
27	323	9.5	109	12	US-10-363-676-6	Sequence 6, Appli
28	280	8.2	139	12	US-10-052-648A-83	Sequence 83, Appli
29	278	8.1	139	12	US-10-052-648A-84	Sequence 84, Appli
30	254	7.4	217	9	US-09-815-419-2	Sequence 2, Appli
31	254	7.4	217	12	US-10-405-808-2	Sequence 2, Appli
32	253.5	7.4	665	9	US-09-816-494-2	Sequence 2, Appli
33	253.5	7.4	665	10	US-09-964-277-2	Sequence 2, Appli
34	252	7.4	171	12	US-10-405-808-21	Sequence 21, Appli
35	247.5	7.3	665	12	US-10-094-749-2312	Sequence 2312, Ap
36	239.5	7.0	444	10	US-09-964-899-47	Sequence 47, Appli
37	239.5	7.0	482	12	US-10-346-356-2	Sequence 2, Appli
38	237.5	7.0	172	9	US-09-816-494-8	Sequence 8, Appli
39	237.5	7.0	172	9	US-09-815-419-5	Sequence 5, Appli
40	237.5	7.0	172	12	US-10-165-272-5	Sequence 5, Appli
41	237.5	7.0	172	12	US-10-410-764-133	Sequence 133, App
42	234	6.9	173	9	US-09-816-494-7	Sequence 7, Appli
43	234	6.9	173	9	US-09-815-419-4	Sequence 4, Appli
44	234	6.9	173	12	US-10-165-272-4	Sequence 4, Appli
45	234	6.9	173	12	US-10-363-676-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-955-732-2
Sequence 2, Application US/09955732
Publication No. US20020182203A1
GENERAL INFORMATION:
APPLICANT: Luche, Ralf M.
TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.433
CURRENT APPLICATION NUMBER: US/09/955,732
CURRENT FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 659
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-732-2

Query Match 100.0%; Score 3412; DB 10; Length 659;
Best Local Similarity 100.0%; Pred. No. 4.2e-263;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MALVTVSRRSPGSGASTPVGPDQAVQRRSRLQRRQSFVLRGAVLGLQDGDNDADA	60
DB	1	MALVTVSRRSPGSGASTPVGPDQAVQRRSRLQRRQSFVLRGAVLGLQDGDNDADA	60
QY	61	SSEPTKAPSEELHGDQTDGQGSQSPQKQEQRQHLMVQLRPQDRIRLAQL	120
DB	61	SSEPTKAPSEELHGDQTDGQGSQSPQKQEQRQHLMVQLRPQDRIRLAQL	120
QY	121	RPRRLRYLLVSTRGEGISQDETLLGVDFPDSSPSCITGLVLPFWMEDTQVYLLDGDG	180
DB	121	RPRRLRYLLVSTRGEGISQDETLLGVDFPDSSPSCITGLVLPFWMEDTQVYLLDGDG	180
QY	181	FSVTSGGQSRIFKPISIQTMWATLQVLHQAACALGSGLVPGSALTWASHYQERLNS	240
DB	181	FSVTSGGQSRIFKPISIQTMWATLQVLHQAACALGSGLVPGSALTWASHYQERLNS	240

QY 241 SCINWTAMADIESLRPPSAEPGGSSSQEQOMEQAIKAEIMKVLVDVSDIESVTSKEIRQAL 300
DB 241 SCINWTAMADIESLRPPSAEPGGSSSQEQOMEQAIKAEIMKVLVDVSDIESVTSKEIRQAL 300
QY 301 ELRLGLPLQOYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQRNVTHTLN 360
DB 301 ELRLGLPLQOYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQRNVTHTLN 360
QY 361 MAREIDNFFPERFTYHNVRMLWDEESAQLLPHWKETHRFIEARAQGTHTLVHCKMGVSR 420
DB 361 MAREIDNFFPERFTYHNVRMLWDEESAQLLPHWKETHRFIEARAQGTHTLVHCKMGVSR 420
QY 421 AATVLAYAMKQYECSLQALRHVQELRPIARPNGFLRQLOIYOGILTASRQSHVWEQKV 480
DB 421 AATVLAYAMKQYECSLQALRHVQELRPIARPNGFLRQLOIYOGILTASRQSHVWEQKV 480
QY 481 GGVSPHEHPAPEVSTPFPPLPPEPEGGEKVKVGMESQAAPKEPGRPRINLRGVMS 540
DB 481 GGVSPHEHPAPEVSTPFPPLPPEPEGGEKVKVGMESQAAPKEPGRPRINLRGVMS 540
QY 541 ISLLEPSLESTSETSDMPEVFSHSHSHEEPLOPFPQARTKGGQYDRGPQALKSR 600
DB 541 ISLLEPSLESTSETSDMPEVFSHSHSHEEPLOPFPQARTKGGQYDRGPQALKSR 600
QY 601 QSVVTLOGSAVVANRTQAFQEQEQGQGGGEPICISSTPRFRKVVRAQSVHDSGEEGEA 659
DB 601 QSVVTLOGSAVVANRTQAFQEQEQGQGGGEPICISSTPRFRKVVRAQSVHDSGEEGEA 659

RESULT 2

US-09-761-640-5
; Sequence 5, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO00964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 5
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Human
US-09-761-640-5

Query Match 99.7%; Score 3402; DB 10; Length 659;
Best Local Similarity 99.8%; Pred. No. 2.6e-262;
Matches 658; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MALVTVSRSPPGSGASTPVGPMDQAVQRRSLQRQSFAYLRGAVLGLQDGDNDDAEA 60
DB 1 MALVTVSRSPPGSGASTPVGPMDQAVQRRSLQRQSFAYLRGAVLGLQDGDNDDAEA 60
QY 61 SSEPTKAPSEELHGDQTDFFGGSSQSPQKQEEQRHLHLMVQLRPODDIRLAQLEAP 120
DB 61 SSEPTKAPSEELHGDQTDFFGGSSQSPQKQEEQRHLHLMVQLRPODDIRLAQLEAP 120
QY 121 RPPRLRYLLVSTREGEGLSQDETVLLGVDFPDSSSPSCTGLVPLMSDTQVYLDGG 180
DB 121 RPPRLRYLLVSTREGEGLSQDETVLLGVDFPDSSSPSCTGLVPLMSDTQVYLDGG 180
QY 161 FSVTSGGQSRIFKPISTQTMWATLQVLHQAEEALGSLVPGSALTWASHYQERLNSQ 240
DB 161 FSVTSGGQSRIFKPISTQTMWATLQVLHQAEEALGSLVPGSALTWASHYQERLNSQ 240
QY 241 SCINWTAMADIESLRPPSAEPGGSSSQEQOMEQAIKAEIMKVLVDVSDIESVTSKEIRQAL 300
DB 241 SCINWTAMADIESLRPPSAEPGGSSSQEQOMEQAIKAEIMKVLVDVSDIESVTSKEIRQAL 300

QY 301 ELRLGLPLQOYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQRNVTHTLN 360
DB 301 ELRLGLPLQOYRDFIDNOMLLVAQRDRASRIFFPHLYLGSEWNAANLEELQRNVTHTLN 360
QY 361 MAREIDNFFPERFTYHNVRMLWDEESAQLLPHWKETHRFIEARAQGTHTLVHCKMGVSR 420
DB 361 MAREIDNFFPERFTYHNVRMLWDEESAQLLPHWKETHRFIEARAQGTHTLVHCKMGVSR 420
QY 421 AATVLAYAMKQYECSLQALRHVQELRPIARPNGFLRQLOIYOGILTASRQSHVWEQKV 480
DB 421 AATVLAYAMKQYECSLQALRHVQELRPIARPNGFLRQLOIYOGILTASRQSHVWEQKV 480
QY 481 GGVSPHEHPAPEVSTPFPPLPPEPEGGEKVKVGMESQAAPKEPGRPRINLRGVMS 540
DB 481 GGVSPHEHPAPEVSTPFPPLPPEPEGGEKVKVGMESQAAPKEPGRPRINLRGVMS 540
QY 541 ISLLEPSLESTSETSDMPEVFSHSHSHEEPLOPFPQARTKGGQYDRGPQALKSR 600
DB 541 ISLLEPSLESTSETSDMPEVFSHSHSHEEPLOPFPQARTKGGQYDRGPQALKSR 600
QY 601 QSVVTLOGSAVVANRTQAFQEQEQGQGGGEPICISSTPRFRKVVRAQSVHDSGEEGEA 659
DB 601 QSVVTLOGSAVVANRTQAFQEQEQGQGGGEPICISSTPRFRKVVRAQSVHDSGEEGEA 659

RESULT 3

US-10-363-676-11
; Sequence 11, Application US/10363676
; Publication No. US20030170856A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN MAP KINASE PHOSPHATASE-LIKE ENZYME
; FILE REFERENCE: LIO122 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/363,676
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: US 60/230,709
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (701)..(701)
; OTHER INFORMATION: xaa-any amino acid
US-10-363-676-11

Query Match 83.9%; Score 2863; DB 12; Length 779;
Best Local Similarity 84.6%; Pred. No. 3e-219;
Matches 581; Conservative 19; Mismatches 39; Indels 48; Gaps 11;
QY 9 SPPGSGASTPV-GP-----WDQAVQRRSLQRQSFAYLR--- 42
DB 24 APVAAAGASTALWGPPIPERGEGTALPALTALGLSRQDRLVQRRSLQRR---ALRCSC 79
QY 43 GAVLGLQDGDNDDAEASSEPTKAPSEELH-GDQTDG-QGSGSPQKQEEQRHLH 100
DB 80 GAVLGLQDGDNDDAEASSEPTKAPSEELHGGTQTDFRCKDPRVRSRRSGSNCL 139
QY 101 WVQLRPQDDIRLAQLEA-----PRPRLRY-LIVSTREGEGLSQDETVL---LGVD 150
DB 140 WV---RRAEAGMTSDLEAQAGTGPLGIRIPCLIVVSTREGR--KSEPRMRSSWDVD 193
QY 151 FPDSSSPSCTGLVPLMSDTQVYLDGSGFSVTSGGQSRIFKPISTQTMWATLQVLHQA 210
DB 194 FPDSSSPSCTGLVPLMSDTQVYLYGDDGFSVTSGGQSRIFKPISTQTMWATLQVLHQA 253
QY 211 CEALGSLVGGGSLTWSHYQERLNSQSCINWTAMADIESLRPPSAEPGGSSSQEQ 270
DB 254 CEALGSLVGGGSLTWSHYQERLNSQSCINWTAMADIESLRPPSAEPGGSSSQEQ 313

QY	271	MEQAIRAELMKVLDVSDLESVTSKETROALELRGLPLQOYRDFIDNOMLLLVAQDRAS	330
Db	314	MERAIRAELMKVLDVGDLSEVASREIROALELRGLPLQOYRDFIDNOMLLLVAQDRAS	373
QY	331	RIFFPHLYLGSEWNAANLEBELORNRVTHILNMAREIDNFYPERFTYHNVRLMDEESAQLLP	390
Db	374	RIFFPHLYLGSEWNAANLEBELORNRVTHILNMAREIDNFYPERFTYHNVRLMDEESAQLLP	433
QY	391	HWKETHRFIEARAQGTHTVLWHCKMGVSRSAATVLAYAMKOYECSEQALRHVOELRPPIA	450
Db	434	HWKETHRFIEARAQGTHTVLWHCKMGVSRSAATVLAYAMKOYECSEQALRHVOELRPPIA	493
QY	451	RPNPGLRQLOIYQGILITASRQSHVMEQKGVSPPEHPAPREVSTPPPLPPEPEGGEE	510
Db	494	RPNPGLRQLOIYQGILITASRQSHVMEQKGVSPPEHPAPREVSTPPPLPPEPEGGEE	553
QY	511	KVVGMEESQAAPKEEPGRPRINLRGVMSISLLEPSLELESTSETSDMPEVFSSHSSH	570
Db	554	KVVGMEESQAAPKEEPGRPRINLRGVMSISLLEPSLELESTSETSDMPEVFSSHSSH	613
QY	571	EEPLQFPQLARTKGGQQVDRGPQPALKSQSVVTLQGSAAVANRTQAFQEQEQGGQGGQ	630
Db	614	EEPLQFPQLARTKGGQQVDRGPQPALKSQSVVTLQGSAAVANRTQAFQEQEQGGQGGQ	673
QY	631	GEPCISSTPRFRKVVRRQASVHDSGEEG	657
Db	674	GEPCISSTPRFRKVVRRQASVHDSGEEG	700

```

RESULT 4
US-10-108-260A-2517
; Sequence 2517, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cdNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2517
;
; LENGTH: 513
;
; TYPE: PRT
;
; ORGANISM: Homo sapiens
;
US-10-108-260A-2517

```

Query Match	76.9%;	Score 2623;	DB 12;	Length 513;
Best Local Similarity	99.0%;	Pred. No. 2.2e-200;		
Matches 506;	Conservative	1;	Mismatches 2;	Indels 2;
				Gaps 1;

QY	151	FP--DSSSPCTGLGLVLP	LMWSDTQVYLYDGDG	FSVTS	GGGSRIFKPI	SIQTMMATLQVYH	208
		:	:	:	:	:	
Db	3	FP	LSPASSPCTGLGLVLP	MSDTQVYLYDGDG	FSVTS	GGGSRIFKPI	SIQTMMATLQVYH 62
QY	209	QACEAALGSGLYPGSALT	WASHYQERLNSEQ	CLNEMWTAMADLES	LRPPSAEPGSSSQ		268
		:	:	:	:	:	
Db	63	QACEAALGSGLYPGSALT	WASHYQERLNSEQ	CLNEMWTAMADLES	LRPPSAEPGSSSQ		122
QY	269	EQMEQAIRAELMKVLD	VSDES	TSKETRQALELRGLP	LOQYRDFIDNQMLLV	QAQRDR	328
		:	:	:	:	:	
Db	123	EQMEQAIRAELMKVLD	VSDES	TSKETRQALELRGLP	LOQYRDFIDNQMLLV	QAQRDR	182
QY	329	ASRIFPHLYLGSEWNAAN	LEELQRNRVTHILNMARE	IDNFPERFTYHNVR	LMWDEESAQL		388
		:	:	:	:	:	
Db	183	ASRIFPHLYLGSEWNAAN	LEELQRNRVTHILNMARE	IDNFPERFTYHNVR	LMWDEESAQL		242
QY	389	LFHWKETHRFIEAARAQ	GTHVLVHCKMGVSRSA	TVLAYAMKQYEC	SLQALRHVQELRP		448
		:	:	:	:	:	
Db	243	LFHWKETHRFIEAARAQ	GTHVLVHCKMGVSRSA	TVLAYAMKQYEC	SLQALRHVQELRP		302
QY	449	IARPNGFLRQLOLYOG	ILTASRQSHWEQKVGV	SPBEHPADBEVSTP	PPPLPPEBEGGG		508

Db	303	IARPNPGLRLQLOLYQGILTLASRQGHWEQKVGVSPHEHBAPEVSTPRLPPEDEGGG	3622
QY	509	BEKVGMEESSQAPAKEEPGPRPRINLRGVMSISLLEPSLEESTSETSDMPEVFSHES	5688
Db	363	GEKVGMEESSQAPAKEEPGPRPRINLRGVMSISLLEPSLEESTSETSDMPEVFSHES	4222
QY	569	SHEEPLQFPQLARTKGGQVDRGQPALKSRQSVVTLQSSAVVANRTOAFQEQEQGQ	6288
Db	423	SHEEPLQFPQLARTKGGQVDRGQPALKSRQSVVTLQSSAVVANRTOAFQEQEQGQ	4882
QY	629	GQGEPCISSTPRFRKVVROASVHDSGEEGEA	6559
Db	483	GQGEPCISSTPRFRKVVROASVHDSGEEGEA	513

```

RESULT 5
US-09-761-640-4
; Sequence 4, Application US/09761640
; Patent No. US20020137042A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000964-CIP
; CURRENT APPLICATION NUMBER: US/09/761,640
; CURRENT FILING DATE: 2001-01-18
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Human
US-09-761-640-4

```

Query Match	71.1%;	Score 2426;	DB 10;	Length 471;
Best Local Similarity	100.0%;	Pred. No. 9.8e-185;		
Matches 469;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MALVTVSRKSPDGS	GASTPVGP	MDQAVOR	KRSRLQRR	ROSFAVL	RGAVLGL	ODG	DNDDAE	60	
Db	1	MALVTVSRKSPDGS	GASTPVGP	MDQAVOR	KRSRLQRR	ROSFAVL	RGAVLGL	ODG	DNDDAE	60	
QY	61	SSEPTKAPSEHEL	GDQDTDFG	QGSQSPQK	QEBOR	OHLMVOLL	RPDDIR	LAAQLE	AP	1200	
Db	61	SSEPTKAPSEHEL	GDQDTDFG	QGSQSPQK	QEBOR	OHLMVOLL	RPDDIR	LAAQLE	AP	1200	
QY	121	RPRLRLYLIVST	REGEGLSQ	DETVLLG	VPDSSPS	CTLGLVLP	PLMSD	TQVYLD	GG	1800	
Db	121	RPRLRLYLIVST	REGEGLSQ	DETVLLG	VPDSSPS	CTLGLVLP	PLMSD	TQVYLD	GG	1800	
QY	181	FSVTSGGQSRIF	KPISIQ	TWATLV	LHQACE	ALGSLV	PGSALT	WASHY	QERLNS	EQ 2400	
Db	181	FSVTSGGQSRIF	KPISIQ	TWATLV	LHQACE	ALGSLV	PGSALT	WASHY	QERLNS	EQ 2400	
QY	241	SCINEMTAMAD	LESRLRPS	APPGSS	EQOMEQ	AIRAEL	MYLV	VS	DLES	VS	TSKEIRQAL 3000
Db	241	SCINEMTAMAD	LESRLRPS	APPGSS	EQOMEQ	AIRAEL	MYLV	VS	DLES	VS	TSKEIRQAL 3000
QY	301	ELRLGLPLQY	RDFIDN	QMLLV	VAQR	DRASRIF	PHLYL	GSE	WNAANLE	ELQ	RNRVTHILN 3600
Db	301	ELRLGLPLQY	RDFIDN	QMLLV	VAQR	DRASRIF	PHLYL	GSE	WNAANLE	ELQ	RNRVTHILN 3600
QY	361	MAREIDNFEY	ERFTYHN	VRWLWDE	ESAOLL	PHWKETH	RFEA	BARA	Q	CTHVLV	HCKMGVRS 4200
Db	361	MAREIDNFEY	ERFTYHN	VRWLWDE	ESAOLL	PHWKETH	RFEA	BARA	Q	CTHVLV	HCKMGVRS 4200
QY	421	AATVLAYAMK	QYEC	SLQAL	RHVQEL	RPIARP	NP	GF	LRLQ	LYQ	GILTA 469
Db	421	AATVLAYAMK	QYEC	SLQAL	RHVQEL	RPIARP	NP	GF	LRLQ	LYQ	GILTA 469

RESULT 6
US-09-955-732-21

Sequence 21, Application US/09955732
 Publication No. US20020182203A1
 GENERAL INFORMATION:
 APPLICANT: Lucbe, Ralf M.
 APPLICANT: Wei, Bo
 TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
 FILE REFERENCE: 200125.433
 CURRENT APPLICATION NUMBER: US/09/955,732
 CURRENT FILING DATE: 2001-09-18
 NUMBER OF SEQ ID NOS: 27
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 21
 LENGTH: 471
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-955-732-21

Query Match 71.1%; Score 2426; DB 10; Length 471;
 Best Local Similarity 100.0%; Pred. No. 9.8e-185;
 Matches 469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSLQRRQSFAYLRGAVLGLQDGGDNDAAEA 60
 DB 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSLQRRQSFAYLRGAVLGLQDGGDNDAAEA 60
 QY 61 SSEPTKAPSEELHGDQTDGFGQSQSPQKQEQRQHLHLMVQLRPQDDIRLAAQLAEP 120
 DB 61 SSEPTKAPSEELHGDQTDGFGQSQSPQKQEQRQHLHLMVQLRPQDDIRLAAQLAEP 120
 QY 121 RPRRLRYLLVSTREGGLSQDETVLLGVDFPDSSPSCTLGLVPLMSDTQVYLDGPG 180
 DB 121 RPRRLRYLLVSTREGGLSQDETVLLGVDFPDSSPSCTLGLVPLMSDTQVYLDGPG 180
 QY 181 FSVTSGGQSRIFKPISTQTMATLQVLHQAACEALGSLVPGSALTWASHYQERLNSQ 240
 DB 181 FSVTSGGQSRIFKPISTQTMATLQVLHQAACEALGSLVPGSALTWASHYQERLNSQ 240
 QY 241 SCLEWTAMADLESIRPSAEPGSSSEQEQEQAIRAEIMKVLVDVDSLESVTSKEIRQAL 300
 DB 241 SCLEWTAMADLESIRPSAEPGSSSEQEQEQAIRAEIMKVLVDVDSLESVTSKEIRQAL 300
 QY 301 ELRLGLPLQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSENNANLELQNRVTHILN 360
 DB 301 ELRLGLPLQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSENNANLELQNRVTHILN 360
 QY 361 MAREIDNFFPERFTYHNVRLMDEESAQLLPHWKETHRFIEAARAQGTHTLVHCKMGVRS 420
 DB 361 MAREIDNFFPERFTYHNVRLMDEESAQLLPHWKETHRFIEAARAQGTHTLVHCKMGVRS 420
 QY 421 AATVLAAYAMKQYECLEQALRHVQELRPIARPNGFLRQLQIYQIGILTA 469
 DB 421 AATVLAAYAMKQYECLEQALRHVQELRPIARPNGFLRQLQIYQIGILTA 469

RESULT 7

US-09-761-640-6
 Sequence 6, Application US/09761640
 Patent No. US20020137042A1
 GENERAL INFORMATION:
 APPLICANT: Wei, Ming-Hui et al
 TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
 TITLE OF INVENTION: AND USES THEREOF
 FILE REFERENCE: CL000964-CIP
 CURRENT APPLICATION NUMBER: US/09/761,640
 CURRENT FILING DATE: 2001-01-18
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 6
 LENGTH: 408
 TYPE: PRT

; ORGANISM: Human
US-09-761-640-6

Query Match 60.1%; Score 2051.5; DB 10; Length 408;
 Best Local Similarity 86.6%; Pred. No. 5.6e-155;
 Matches 406; Conservative 0; Mismatches 0; Indels 63; Gaps 1;

QY 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSLQRRQSFAYLRGAVLGLQDGGDNDAAEA 60
 DB 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSLQRRQSFAYLRGAVLGLQDGGDNDAAEA 60
 QY 61 SSEPTKAPSEELHGDQTDGFGQSQSPQKQEQRQHLHLMVQLRPQDDIRLAAQLAEP 120
 DB 61 SSEPTKAPSEELHGDQTDGFGQSQSPQKQEQRQHLHLMVQLRPQDDIRLAAQLAEP 120
 QY 121 RPRRLRYLLVSTREGGLSQDETVLLGVDFPDSSPSCTLGLVPLMSDTQVYLDGPG 180
 DB 121 RPRRLRYLLVSTREGGLSQDETVLLGVDFPDSSPSCTLGLVPLMSDTQVYLDGPG 180
 QY 181 FSVTSGGQSRIFKPISTQTMATLQVLHQAACEALGSLVPGSALTWASHYQERLNSQ 240
 DB 181 FSVTSGGQSRIFKPISTQTMATLQVLHQAACEALGSLVPGSALTWASHYQERLNSQ 240
 QY 241 SCLEWTAMADLESIRPSAEPGSSSEQEQEQAIRAEIMKVLVDVDSLESVTSKEIRQAL 300
 DB 241 SCLEWTAMADLESIRPSAEPGSSSEQEQEQAIRAEIMKVLVDVDSLESVTSKEIRQAL 300
 QY 301 ELRLGLPLQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSENNANLELQNRVTHILN 360
 DB 301 ELRLGLPLQYRDFIDNOMLLVAQRDRASRIFFPHLYLGSENNANLELQNRVTHILN 360
 QY 361 MAREIDNFFPERFTYHNVRLMDEESAQLLPHWKETHRFIEAARAQGTHTLVHCKMGVRS 420
 DB 361 MAREIDNFFPERFTYHNVRLMDEESAQLLPHWKETHRFIEAARAQGTHTLVHCKMGVRS 420
 QY 421 AATVLAAYAMKQYECLEQALRHVQELRPIARPNGFLRQLQIYQIGILTA 469
 DB 421 AATVLAAYAMKQYECLEQALRHVQELRPIARPNGFLRQLQIYQIGILTA 469

RESULT 8

US-09-761-640-8
 Sequence 8, Application US/09761640
 Patent No. US20020137042A1
 GENERAL INFORMATION:
 APPLICANT: Wei, Ming-Hui et al
 TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
 TITLE OF INVENTION: AND USES THEREOF
 FILE REFERENCE: CL000964-CIP
 CURRENT APPLICATION NUMBER: US/09/761,640
 CURRENT FILING DATE: 2001-01-18
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 8
 LENGTH: 312
 TYPE: PRT
 ORGANISM: Human
 US-09-761-640-8

Query Match 43.1%; Score 1470.5; DB 10; Length 312;
 Best Local Similarity 92.9%; Pred. No. 7.6e-109;
 Matches 290; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

QY 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSLQRRQSFAYLRGAVLGLQDGGDNDAAEA 60
 DB 1 MALVTVSRSPPGSGASTPVGPWDQAVQRRSLQRRQSFAYLRGAVLGLQDGGDNDAAEA 60
 QY 61 SSEPTKAPSEELHGDQTDGFGQSQSPQKQEQRQHLHLMVQLRPQDDIRLAAQLAEP 120
 DB 61 SSEPTKAPSEELHGDQTDGFGQSQSPQKQEQRQHLHLMVQLRPQDDIRLAAQLAEP 120
 QY 121 RPRRLRYLLVSTREGGLSQDETVLLGVDFPDSSPSCTLGLVPLMSDTQVYLDGPG 180

121 RPPRLRYLLVSTREGGLSODETVLLGVDFPDSSSPCTGLVLPWMSDTQVYLDGDDG 180
QY 181 FSVTSGGQSRIFKPISIQTMWATLQVLAQCEALGSLVPGSALTWASHYQERLNSQ 240
Db 181 FSVTSGGQSRIFKPISIQTMWATLQVLAQCEALGSLVPGSALTWASHYQERLNSQ 240
QY 241 SCLNWTAMADLESIRPPSAEPGSGSEQEQAIRAEIMKVLVDV-SDLESYTSKEIRQA 299
Db 241 SCLNWTAMADLESIRPPSAEPGSGSEQEQAIRAEIMKVLVELESTSETSDMPEVFS 300
QY 300 LELRLGLPLQY 311
Db 301 HESSHEEPLQPF 312

RESULT 9
US-10-108-260A-3142
; Sequence 3142, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3142
; LENGTH: 703
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3142

Query Match 33.3%; Score 1136; DB 12; Length 703;
Best Local Similarity 42.9%; Pred. No. 1.1e-81;
Matches 256; Conservative 91; Mismatches 152; Indels 98; Gaps 17;
QY 36 QSFAYLRGAVLGLQDGGDNDAAEASSEPTTEKAPSEBELHGDQTFGQGS-QSPQOE-E 93
Db 51 ESFFMVKGALFLQOG-----SSPQ-----GQSLQHPHKGAGD 84
QY 94 QROHLHLMVQLRPQDDIRLAQLEAPRPRRLRYLLV--STREGGLSODETVLLGVDF 151
Db 85 LPQHLQVMINLRCEDRIKLAVRLESAMADVRYMVVYSSGRQ---DTEENILGVDF 140
QY 152 PDSSSPSCTGLVLPWMSDTQVYLDGDDGFSVTSGGQSRIFKPISIQTMWATLQVLAQ 211
Db 141 SSKESSCTIGWLRIMSDTKIHLDDGGFSVSTAGRMIKFPVSVQAMWSALQVLAQAC 200
QY 212 EALGSLVPGSALTWASHYQERLNSQSCINWTAMADLESIRP--PSAEPGSGSEQE 269
Db 201 EVARRHNYFPGVALIWAITYEESCISSEQSCINENAMQDLESTRPDSPALFVDKPTGE 260
QY 270 QMEQAIRAEIMKVLVDVSDLESYTSKEIRQALEIRLGLPLQYRDFIDNQMLLVAQRDA 329
Db 261 RTERLIKAKLRISIMSQDLENVTSKEIRNELEKQMCNMLKEKLFIDNEMILLQMDKP 320
QY 330 SRIFPHLYLGESEWNAANLEELQNRVTHILNMAREIDNFFPERFTYHNVRIMDEESAQL 389
Db 321 SLIFDHLYLGESEWNASLUELQSGVDYILNVTREIDNFFPGLFAVHNIRVYDEETDIL 380
QY 390 PHMKETHRFIEAARAQSTHVLVHCKMGVSRSAATVLAAYAMKQYEGSLQALRHVQELRPI 449
Db 381 AHMNEAYHFIKAKRNSKCLVHCKMGVSRSASTVIAYAMKEFGWPLEKAYNYVKQKSI 440
QY 450 ARPNGFLRQLQIYQGITLASRQSH--VMEQKVGG--VSPHEHPA-----PE 492
Db 441 TRPNAGFMRQLSEYEGILDASKQRHNKLMWRQQTDSLQQPVDDPAGPDFLPETPDGTPE 500
QY 493 VSTPF-----PPLP-----PEPEGGEKVVGMEE-----SQAP 522
Db 501 SQLPFLDDAAQPGGLPPLPCCFRRRLSDPLSPEDETG---SLVHLEDPEREALEAP 557

QY 523 KEE---PGPRPRINRGVMRSISLLEPSL--ELESTSETSDMPEVFSHESHEEPL 574
Db 558 PAEVHRPARQPQ-----QGSGLCEKDYKKLEFGSPKGRGSGLLQVEETEREEGL 607

RESULT 10
US-09-775-925-6
; Sequence 6, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775, 925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-775-925-6

Query Match 31.0%; Score 1058; DB 9; Length 509;
Best Local Similarity 42.5%; Pred. No. 1.2e-75;
Matches 213; Conservative 100; Mismatches 134; Indels 54; Gaps 7;
QY 36 QSFAYLRGAVLGLQDGGDNDAAEASSEPTTEKAPSEBELHGDQTFGQGSQSPQOEQ- 94
Db 46 ESFLTVKGALFLPR-----GNGSSTPRISHRN 74
QY 95 -----ROHLHLMVQLRPQDDIRLAQLEAPRPRRLRYLLVSTREGGLSODETVLLG 148
Db 75 KHAGDLQQLQAMFILRPEDNIRLAVRLESTYQNRTRYMVVSTNGRQ--DTEESIVLG 132
QY 149 VDFPDSSSPSCTGLVLPWMSDTQVYLDGDDGFSVTSGGQSRIFKPISIQTMWATLQVLA 208
Db 133 MDFSNDSSCTGMGLVLPWMSDTLHLDDGGFSVSTDNRVHIFKPVSVQAMWSALQSLH 192
QY 209 QACEAALGSLVPGSALTWASHYQERLNSQSCINWTAMADLESIRP--PSAEPGSGSS 266
Db 193 KACEYARAHHNYFPGSLFLTWVSYESHINSQSSVNEENAMQDVQSHRPDSPALFTDIFT 252
QY 267 EQEQAIRAEIMKVLVDVSDLESYTSKEIRQALEIRLGLPLQYRDFIDNQMLLVAQR 326
Db 253 ERERTERLIKTKLRREIMQDLENITSKEIRLEEMQMCNMLREFKEFIDNEMIVILQM 312
QY 327 DRASRIFPHLYLGESEWNAANLEELQNRVTHILNMAREIDNFFPERFTYHNVRIMDEESA 386
Db 313 DSPIDFHEHVLGESEWNASLLEDLQNRGVRYILNVTREIDNFFPGVEYHNIRVYDEEAT 372
QY 387 QLEPHMKETHRFIEAARAQSTHVLVHCKMGVSRSAATVLAAYAMKQYEGSLQALRHVQEL 446
Db 373 DLLAYWMDTYKFIKAKRGSKCLVHCKMGVSRSASTVIAYAMKEYNLDRAYDYKER 432
QY 447 RPIARNPGFLRQLQIYQGITLASRQSHVMEQKVGVSPHEHPAPEVSTPFP-----PPLP 501
Db 433 RIVTKPNPSFMRQLEEYQGITLASFLGLIH---GG---RDKPWEKSTEFESVDLVISIP 485
QY 502 PEPEGGEKVVGMEEQAAP 522
Db 486 GSPSCNPBKLLHISHPYLTP 506

RESULT 11
US-09-955-732-14
; Sequence 14, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Lucche, Ralf M.

```

; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732
; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-955-732-14

Query Match      31.0%; Score 1058; DB 10; Length 509;
Best Local Similarity 42.5%; Pred. No. 1.2e-75;
Matches 213; Conservative 100; Mismatches 134; Indels 54; Gaps 7;

QY 36 QSFAYLRGAVLGLQDGGDNDAAEASSEPTEKAPSEELHGDQDTFGQSGSPQKQEQ- 94
DB 46 ESFLTVKGAALFLPR-----GNGSSTPRISHRRN 74
QY 95 -----RQHLHLMVQLRPQDDIRLAQLEAPRPRLRYLLVSTREGELSDQETVLG 148
DB 75 KHAGDIQHLQAMFILLRPEDNIRLAVRLESTYQNRTRYMVVSTNGRQ--DTEESIVLG 132
QY 149 VDFPDSSPSCTGLVLPMSDTQVYLDGSGSVTSYGQSRIFKPSIQTMATLQVLRH 208
DB 133 MDFSNDSSCTGTGLVLPMSDTLHLDGSGSVSTDNRVHIFKPSVQAMWSALQSLH 192
QY 209 QACEAALGSLVPGSALTWASHYQERLNEQSCLENWTAMADLESRLP--PSAEPGSS 266
DB 193 KACEVARAHNYYPGSLFTWVSYESHINSQSVNEMAMQDVQSHRPDSPALFTDPT 252
QY 267 EQQMEQAIRAELMKVLDVSDLESVTSKEIRQALELRLGLPLQYRDFIDNQLLVAGR 326
DB 253 ERERTERLIKTKLRIMWQDLENITSKEIRTELEMQMCNLRFEKEFIDNEMIVILGQM 312
QY 327 DRASRIFPHLYLGSEWNAANLELQNRVTHILNMAREIDNFYPERFTYHNVRMLDESA 386
DB 313 DSPTOIFEHVFLGSEWNASLIEDLQNRGVRYILNVTREIDNFFPGVFEYHNIRVYDEAT 372
QY 387 QLPWKETHRFIEARAGCTHVLVHCKMGVSRSAATVLAAMKQVECSLEQALRHVQEL 446
DB 373 DLLAYMNDTYKFSKAKKGSKCLVHCKMGVSRSASTVIAYAMKEYGNLDRAYDYKER 432
QY 447 RPIARPNGFLRQLOIYQGLLTASRQSHWQKVGVSPEHPAPEVSTPFP-----PLP 501
DB 433 RTVTKNPNSFMKQLEBYQGLLASFLGLIH---CG--RDKPWGKSTEFESVDLVSLP 485
QY 502 PEPEGGEKVVGMESQAP 522
DB 486 GSPSCCNPEKLLHISHPYLTP 506

RESULT 12
US-10-181-590-9
; Sequence 9, Application US/10181590
; Publication No. US20030152949A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: MATHUR, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: LAL, Preeti
; APPLICANT: WANG, Eureka
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: POLICKY, Jennifer L.
; TITLE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: PI-0018 PCT
```

```

; CURRENT APPLICATION NUMBER: US/10/181,590
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 60/177,719; 60/178,988; 60/184,959; 60/190,142
; PRIOR FILING DATE: 2000-01-21; 2000-01-28; 2000-02-25; 2000-03-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030152949A1 637331CD1
; US-10-181-590-9

Query Match      29.5%; Score 1005; DB 12; Length 484;
Best Local Similarity 40.7%; Pred. No. 1.8e-71;
Matches 211; Conservative 97; Mismatches 143; Indels 68; Gaps 8;

QY 1 MALTVSRSPPGSGASTPVGEWDQAVQRRSLQRR--QSFAYLRGAVLGLQDGGDNDAA 57
DB 1 MALTVQRSPTPSTSSPCASEADSGEEBCRSQPRSSISESLTVKGAALFLPR----- 53
QY 58 AEASSEPTEKAPSEEBELHGDQDTFGQSGSPQKQEQ-----RQHLHLMVQLRPQDD 110
DB 54 -----GNGSSTPRISHRRNKHAGDIQHLQAMFILLRPEDN 89
QY 111 IRLAQLEAPRPRLRYLLVSTREGELSDQETVLGVDFPDSSPSCTGLVLPMSD 170
DB 90 IRLAVRLESTYQNRTRYMVVSTNGRQ--DTEESIVLGMDFSNDSSCTGTGLVLPMSD 147
QY 171 TQVYLDGSGFSVTSYGQSRIFKPSIQTMATLQVLRHQAACEAALGSLVPGSALTWAS 230
DB 148 TLHLDGSGFSVSTDNRVHIFKPSVQAMWSALQSLHAKACEVARAHNYYPGSLFTWVS 207
QY 231 HYQERLNEQSCLENWTAMADLESRLP--PSAEPGSSSEQQMEQAIRAELMKVLDVSD 288
DB 208 YYESHINSQSVNEMAMQDVQSHRPDSPALFTDPTERERTERLIKTKLRIMWQD 267
QY 289 ESVTSKEIRQALELRLGLPLQYRDFIDNQLLVAGRDRASRIFPHLYLGSEWNAANLE 348
DB 268 ENITSKEIRTELEMQMCNLRFEKEFIDNEMIVILGQMDSPTOIFEHVFLGSEWNASL 327
QY 349 ELQNRVTHILNMAREIDNFYPERFTYHNVRMLDESAQLLPWKETHRFIEARAGCTH 408
DB 328 DLQNRGVRYILNVTREIDNFFPGVFEYHNIRVYDEATDLLAYMNDTYKFSKAKKGS 387
QY 409 VLVHCKMGVSRSAATVLAAMKQVECSLEQ-----LRHVQELRPIARPNGFLRQLOIY 463
DB 388 CLVHCKMGVSRSASTVIAYAMKEYGNLDRASSRIRLEKVTNLMGF-----DVCTF 438
QY 464 QGILTASRQSHWQKVGVSPEHPAPEVSTPFPPLP 502
DB 439 PTIMTI-----QIMNISITPKKFPYASISFSPLSP 468

RESULT 13
US-09-775-925-2
; Sequence 2, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucite, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
```

ORGANISM: Homo sapiens
US-09-775-925-2
Query Match 28.5%; Score 971; DB 9; Length 552;
Best Local Similarity 44.6%; Pred. No. 1.1e-68;
Matches 208; Conservative 73; Mismatches 121; Indels 64; Gaps 11;
QY 163 LVLPWMSDTQVYLDGSGFSVTSGGQSRIFKPISTQTMATLQVHLHQAACEALGSLVPG 222
Db 1 MVLRLMSDTKIHLDGSGFSVSTAGRMIIFKPVSVQAMMSALQVHLKACEVARRHNYFPG 60
QY 223 GSALTWASHYQERLNSEQSCINENWTAMADLESRLP--PSAEPGSSSEQEQMEQATRAELW 280
Db 61 GVALIMATYYESCTISSEQSCINENWAMQDLESTRPDSPALFVDKPTGERTERLIKAKLR 120
QY 281 KVLVDSDLESVTSKEIRQALERLGLPQQYRDFIDNQMLLVAQRDRASRIFFHLYLGS 340
Db 121 SIMSQDLENVTSKEIRNELKQMCNLIKELKEFIDNEMLLIGQMDKPSLIFDHYLGS 180
QY 341 EWNANLEELQNRVTHILNMAREIDNFFPERFTYHNRLWDEESAQLPHWKETHRFIE 400
Db 181 EWNASNLEELQSGVDYILNVTREIDNFFPGLFAYHNIRVYDETTDLLAHMNEAYHFIN 240
QY 401 AARAQGTHTLVHCKMGVSRSAATVLAYAMKQYECLEQALRHVQELRPIARNPGFLRQL 460
Db 241 KAKRNHCLKVHCKMGVSRSASTVIAYAMKEFGWPLEKAYNYVKQKRSITRNAGFMROL 300
QY 461 QIYQGITLTSRQSH--VWEQKVG--VSPHEHPA-----PEVSTPF----- 497
Db 301 SEYEGILDASKQRHNKLMRQQTDSLQQPVDDPAGPGDFLPETPDGTBPESQLPFLDDAQ 360
QY 498 ----PPLP-----PEPEGGEKVGVMEE-----SQAAPKEE--PGPRP 530
Db 361 PGLGPPLPCCFRRRLSDPLLPSPEDETG--SLVHLEDPEREALLBEAAPPAEVHRRPARQP 417
QY 531 RINLRGVMRSISLLEPSL--ELESTSETSDMPEVFSHESHESHEEPL 574
Db 418 Q-----QSGGLCEKDVKKLEFGSPKGRSGSLLOVEETEREEGL 456

RESULT 14
US-09-955-732-15
Sequence 15, Application US/09955732
Publication No. US20020182203A1
GENERAL INFORMATION:
APPLICANT: Lucche, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.433
CURRENT APPLICATION NUMBER: US/09/955,732
CURRENT FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 552
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-732-15

Query Match 28.5%; Score 971; DB 10; Length 552;
Best Local Similarity 44.6%; Pred. No. 1.1e-68;
Matches 208; Conservative 73; Mismatches 121; Indels 64; Gaps 11;
QY 163 LVLPWMSDTQVYLDGSGFSVTSGGQSRIFKPISTQTMATLQVHLHQAACEALGSLVPG 222
Db 1 MVLRLMSDTKIHLDGSGFSVSTAGRMIIFKPVSVQAMMSALQVHLKACEVARRHNYFPG 60
QY 223 GSALTWASHYQERLNSEQSCINENWTAMADLESRLP--PSAEPGSSSEQEQMEQATRAELW 280
Db 61 GVALIMATYYESCTISSEQSCINENWAMQDLESTRPDSPALFVDKPTGERTERLIKAKLR 120
QY 281 KVLVDSDLESVTSKEIRQALERLGLPQQYRDFIDNQMLLVAQRDRASRIFFHLYLGS 340

Db 121 SIMSQDLENVTSKEIRNELKQMCNLIKELKEFIDNEMLLIGQMDKPSLIFDHYLGS 180
QY 341 EWNANLEELQNRVTHILNMAREIDNFFPERFTYHNRLWDEESAQLPHWKETHRFIE 400
Db 181 EWNASNLEELQSGVDYILNVTREIDNFFPGLFAYHNIRVYDEETDLLAHMNEAYHFIN 240
QY 401 AARAQGTHTLVHCKMGVSRSAATVLAYAMKQYECLEQALRHVQELRPIARNPGFLRQL 460
Db 241 KAKRNHCLKVHCKMGVSRSASTVIAYAMKEFGWPLEKAYNYVKQKRSITRNAGFMROL 300
QY 461 QIYQGITLTSRQSH--VWEQKVG--VSPHEHPA-----PEVSTPF----- 497
Db 301 SEYEGILDASKQRHNKLMRQQTDSLQQPVDDPAGPGDFLPETPDGTBPESQLPFLDDAQ 360
QY 498 ----PPLP-----PEPEGGEKVGVMEE-----SQAAPKEE--PGPRP 530
Db 361 PGLGPPLPCCFRRRLSDPLLPSPEDETG--SLVHLEDPEREALLBEAAPPAEVHRRPARQP 417
QY 531 RINLRGVMRSISLLEPSL--ELESTSETSDMPEVFSHESHESHEEPL 574
Db 418 Q-----QSGGLCEKDVKKLEFGSPKGRSGSLLOVEETEREEGL 456

RESULT 15
US-09-761-640-9
Sequence 9, Application US/09761640
Patent No. US20020137042A1
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
TITLE OF INVENTION: AND USES THEREOF
FILE REFERENCE: CL000964-CIP
CURRENT APPLICATION NUMBER: US/09/761,640
CURRENT FILING DATE: 2001-01-18
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 524
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-761-640-9

Query Match 26.9%; Score 918.5; DB 10; Length 524;
Best Local Similarity 37.9%; Pred. No. 1.6e-64;
Matches 205; Conservative 83; Mismatches 164; Indels 89; Gaps 7;
QY 1 MALVTYVRSPPGSGA-STPYGPMQAVQRRSLQRRQSFVLRGAVLGLQDGDNDDAE 59
Db 1 MALVTYVRSPPSVAGSCNSDGESEDDGNSKGNDRSECFAGKGTALV----- 49
QY 60 ASSEPTKAPSEELHGDQTFQGGSQSPQKQEQRQHLHLMVQLRPQDRIQAQLEA 119
Db 50 ALKDIPPLTQSERRLSTDSTRSSNSTQS--NNSDQLHLQSMFYLLQREDTLKMAVKLES 107
QY 120 PRPPRLRYLLVS----- 132
Db 108 QRSNRTYLVIASRSCCRSGTSDRRHRIMRHHSVYVGSAGTKSSTSPAVPTQRLSVE 167
QY 133 -----TREGGLS---QDETVLIGVDFPDDSSPSCTLGLVPLWSD 170
Db 168 QTATEASSCKDKTADKENATAGDNKNTSGMEESCLIGIDCNERT-----TIGLVVPILAD 223
QY 171 TQVYLDGSGFSVTSGGQSRIFKPISTQTMATLQVHLHQAACEALGSLVPGSALTWAS 230
Db 224 TTIHLDGSGFSVKYKETHIFKPVSVQAMMSALQTLHKVSKKARENNFYASGPSHDWLS 283
QY 231 HYQERLNSEQSCINENWTAMADLESRLPSPAEP--GSSSEQEQMEQATRAELWVLDVSD 286
Db 284 SYERRIESDQSCINENWAMDALESRRPPSPDAIRNKPPEKETESVIMKLIKAIMSVDL 343
QY 289 ESVTSKEIRQALERLGLPQQYRDFIDNQMLLVAQRDRASRIFFHLYLGS EWNANLE 348

Db	344	DEVTSKYIRGRLEBILDMDLGEYKSFIDAEMLVILGOMDAPTKIFEHVILGSEWNASNLE	403
QY	349	ELORNRVTHILNMAREIDNPFPERFTYHVRWLWDEESAQLPHWKETHRFIEARAQGT	408
Db	404	ELQKNGVRIILNVTREIDNFPFGTFEYFNVRVYDDEKTNLLKYMDDTFRIYITRAKAEGSK	463
QY	409	VLVHCKMGVSRSAATVLAAMKQYECSEALRHVQELRPIARPNGFLRQLOIYQGI	468
Db	464	VLVHCKMGVSRSAASVIAAMKAYQWEPQALEHVKKRSCIKPNKNFLNQLETYSGMLD	523
QY	469	A 469	
Db	524	A 524	

Search completed: January 15, 2004, 06:55:08
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 06:49:57 : Search time 21 Seconds
(without alignments)
1327.755 Million cell updates/sec

Title: US-09-955-732-2
Perfect score: 3412
Sequence: 1 MALVTVSRSPPGSGASTPVG.....RRKVVROASVHDSGEGEA 659

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	239.5	7.0	482	US-09-557-921-2	Sequence 2, Appli
2	237.5	7.0	172	US-09-704-139-5	Sequence 5, Appli
3	234	6.9	173	US-09-704-139-4	Sequence 4, Appli
4	232.5	6.8	397	US-08-990-379-8	Sequence 8, Appli
5	219	6.4	170	US-09-557-921-14	Sequence 14, Appli
6	219	6.4	171	US-09-544-716-18	Sequence 18, Appli
7	219	6.4	171	US-09-557-921-19	Sequence 19, Appli
8	218.5	6.4	1784	US-09-770-595A-3	Sequence 3, Appli
9	215.5	6.3	176	US-09-704-139-2	Sequence 2, Appli
10	215.5	6.3	395	US-08-990-379-5	Sequence 5, Appli
11	214	6.3	211	US-09-544-716-2	Sequence 2, Appli
12	214	6.3	226	US-09-045-973-8	Sequence 8, Appli
13	213.5	6.3	174	US-09-544-716-20	Sequence 20, Appli
14	211.5	6.2	170	US-09-544-716-14	Sequence 14, Appli
15	211.5	6.2	170	US-09-557-921-15	Sequence 15, Appli
16	211	6.2	207	US-09-013-881-4	Sequence 4, Appli
17	211	6.2	207	US-09-612-473-4	Sequence 4, Appli
18	209.5	6.1	393	US-08-990-379-4	Sequence 4, Appli
19	209	6.1	168	US-09-544-716-15	Sequence 15, Appli
20	209	6.1	168	US-09-557-921-16	Sequence 16, Appli
21	209	6.1	313	US-08-990-379-7	Sequence 7, Appli
22	209	6.1	314	US-09-164-193-22	Sequence 22, Appli
23	209	6.1	314	US-09-221-448A-22	Sequence 22, Appli
24	206.5	6.1	394	US-08-530-290-23	Sequence 23, Appli
25	206.5	6.1	394	US-09-702-705-805	Sequence 805, App
26	206.5	6.1	394	US-09-736-457-805	Sequence 805, App
27	206	6.0	661	US-09-770-595A-22	Sequence 22, Appli

28	205.5	6.0	394	4	US-09-702-705-827	Sequence 827, App
29	205.5	6.0	394	4	US-09-736-457-827	Sequence 827, App
30	204.5	6.0	302	4	US-09-702-705-806	Sequence 806, App
31	204.5	6.0	302	4	US-09-736-457-806	Sequence 806, App
32	203	5.9	169	4	US-09-544-716-17	Sequence 17, Appli
33	203	5.9	169	4	US-09-557-921-18	Sequence 18, Appli
34	202	5.9	169	4	US-09-544-716-16	Sequence 16, Appli
35	202	5.9	169	4	US-09-557-921-17	Sequence 17, Appli
36	202	5.9	367	2	US-08-530-290-24	Sequence 24, Appli
37	200	5.9	263	3	US-09-164-193-5	Sequence 5, Appli
38	200	5.9	263	3	US-09-221-448A-5	Sequence 5, Appli
39	198.5	5.8	223	4	US-09-685-853A-2	Sequence 2, Appli
40	196	5.7	87	4	US-09-371-671B-4	Sequence 4, Appli
41	194	5.7	314	4	US-09-371-671B-11	Sequence 11, Appli
42	194	5.7	367	2	US-08-990-379-6	Sequence 6, Appli
43	193	5.7	168	4	US-09-544-716-13	Sequence 13, Appli
44	193	5.7	168	4	US-09-557-921-13	Sequence 13, Appli
45	191.5	5.6	353	3	US-09-013-881-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-557-921-2
Sequence 2, Application US/09557921
Patent No. 6551810
GENERAL INFORMATION:
APPLICANT: Luche, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.416
CURRENT APPLICATION NUMBER: US/09/557, 921
CURRENT FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 482
TYPE: PRT
ORGANISM: Homo sapiens
US-09-557-921-2

Query Match 7.0%; Score 239.5; DB 4; Length 482;
Best Local Similarity 28.0%; Pred. No. 2.5e-13;
Matches 73; Conservative 44; Mismatches 117; Indels 27; Gaps 6;

QY	232	YQERLNSQSCLENTAMADLESRLPPSAEP----	GGSSQEQMEQAI---RAELWKVLD	284
DB	236	YDENTNEPSRVMPSQPLHIVLESLEKREKEDLVLKGLSFKQNHENLQDMSLOQECRE		295
QY	285	VSDESVTSKEIRQALELRGLPLQQRDFIDNQMLLVAAQRDRASRIFFHLYLGEWNA		344
DB	296	VGGGSAASSLPQ-----PIPTPD-LENAEL-----TPIPLFLGNEQDA		337
QY	345	ANLEELQNRVTHILNMAREIDNFYPER--FTYHNVRLMDEESAQLPHWKETHRFIEAA		402
DB	338	QDLDTMQRNLNGYVINTVTHLLPLHYEKGLEFNKRLPATDSNKQNLQYFEFAFEFIEEA		397
QY	403	RAQGTIVLVHCKMGVSRAATVLAAYAMKQYECLEQALRHVQELRPIARPNGFLRLQI		462
DB	398	HQCKGILLIHQAGVSRSAITIVIALMKHTMTMTDAYKFKGKRPIISPLNFMGQLLE		457
QY	463	YQGILTASRQSHWEQKVGCV	483	
DB	458	FEEDLNGVTPIRLTPKLMGV	478	

RESULT 2
US-09-704-139-5
Sequence 5, Application US/09704139
Patent No. 6420153
GENERAL INFORMATION:
APPLICANT: Kapeller, Rosanna


```

: TITLE OF INVENTION: 16232. A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
: FILE REFERENCE: 10448-018001
: CURRENT APPLICATION NUMBER: US/09/704,139
: CURRENT FILING DATE: 2000-11-01
: PRIOR APPLICATION NUMBER: US 60/185,772
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 5
: LENGTH: 172
: TYPE: PRT
: ORGANISM: Artificial/Unknown
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)..(172)
: OTHER INFORMATION: consensus sequence
US-09-704-139-5

```

```

Query Match      7.0%; Score 237.5; DB 4; Length 172;
Best Local Similarity 37.9%; Pred. No. 6.8e-14;
Matches 64; Conservative 23; Mismatches 47; Indels 35; Gaps 7,

QY      330  SRI
```

```

RESULT 3
US-09-704-139-4
; Sequence 4, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(173)
; OTHER INFORMATION: consensus sequence
US-09-704-139-4

```

```

Query Match          6.9%; Score 234; DB 4; Length 173;
Best Local Similarity 35.9%; Pred. No. 1.4e-13;
Matches 61; Conservative 26; Mismatches 47; Indels 36; Gaps 6,
QY      330 SRIFPHLYLGSEWNA--ANL'EELQRNRVTHLNNAREIDNFYP----- 370
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      3  SEILPHLYLGSYSTASEANLALLKLGITHIVINTEVEVNPFLDKKNDRHYNAYISKN 62
QY      371 ERFTYHNVR-----LWDEESAQLLPHMKETHRFIEARAQGIHLVHCKMGVSR 420
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      63 SGFTYLIPIPNVDHLYYHIAMNHET-KISKYFBEAVDFIDARQKGKVLVHCQAGISRS 121
QY      421 AATVLAYAMKQYECSLEQA-----LRHVQELR-PIARPNGFLRLQLTQ 464
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db          122 ATLIIAYLMKTRNLSLNEAYDFVYVYHIKERRCPIISPNFGFLRQLEIYE 171

RESULT 4
US-08-990-379-8
; Sequence 8, Application us/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Philip J
; APPLICANT: Misra-Press, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; TITLE OF INVENTION: Their Biologically Active Expression Products
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-990-379-8

```

	Query Match	6.8%;	Score 232.5;	DB 2;	Length 397;	
	Best Local Similarity	27.8%;	Pred. No. 7.8e-13;			
	Matches 74;	Conservative 49;	Mismatches 106;	Indels 37;	Gaps 9	
QY	287 DLESVTSKEI--RQALELRGLGFL--QQYRDFFIDNQMLLLVAQRDRASRIFFPHLYLGSEW	342				
Dd	141 DVKPIISQEKIESERALISQCGKPVVNSYRPAYD-----QGGPVEILLPPLYIGSAY	191				
QY	343 NAANIEELQRRNVTHILNMAREIDNFPYPERFTYHNVRWLDEESAQLLPHWKETHRFIBAA	402				
Dd	192 HASKCEPLANLHITALINVSRTSEACMTHLHYKWIPVEDSHADISSHFQEAIIDFIDCV	251				
QY	403 RAQGTHTVLVCKMGVSRSAATVLAAYAMKQYECSLEQA LRHVQELRLPIARPNGFLRQIQI	462				
Dd	252 REKGGRYLWHCEAGISRSPFCIMAYLMKTOKQRLKEAFDYIKQRSMVSPNFEGFMGLLO	311				
QY	463 YOGIL---TASRQSHVWEQKYGVGSPEEH---PAPEVS---TPFRP-----LPPEDEGG	507				
Dd	312 YESEILPSTPNPQPSCQGEAAGSSLI GHLQTLSFDMDGAYCTFPASVLA RCLPTQOSQS	371				
QY	508 GEEKV-----VGMEESQAAPKEE	525				
Dd	372 SAEALWQRPNPAKTGMBES-AQPFQEQ	396				

```

RESULT 5
US-09-557-921-14
; Sequence 14, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557, 921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-14

Query Match      6.4%; Score 219; DB 4; Length 170;
Best Local Similarity 35.1%; Pred. No. 3.4e-12;

```

	Matches	46;	Conservative	28;	Mismatches	55;	Indels	2;	Gaps	1,
QY	332	IFPHLYLCSMNNAANLEELQNRVTHILNMAREIDNEYPER--FTYHNVRLMDEBSAQL								389
Db	30	ILPFLFLGNEQDAQDLDTMQRLNIGVINVTTHLPLHYEKGFLFNFKRLPATDSNKNQNR								89
QY	390	PHMKETHRFIEARAQGTHTVLVHCKMGVSRSAATVLAAMKQYECSLQALRHVQELRPI								449
Db	90	QYFEAEAFEFIEAHQCGKGLIHQAGVSRSAATVLAALMKHTRMTMTDAYKFKGKRPI								149
QY	450	ARPNGFLRQL								460
Db	150	ISPNNLMGQL								160

```

RESULT 6
US-09-544-716-18
; Sequence 18, Application US/09544716
; Patent No. 6492157
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.415
; CURRENT APPLICATION NUMBER: US/09/544,716
; CURRENT FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 18
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-716-18

```

	Query Match	6.4%;	Score 219;	DB 4;	Length 171;
	Best Local Similarity	33.8%;	Pred. No. 3.4e-12;		
	Matches 47;	Conservative 29;	Mismatches 63;	Indels 0;	Gaps 0
QY	332 IFPHLYIGSEWNAANLBELQRNRVTHILNMAREIDNFEYPERFTYHNVRWLWDEESAQLLPH	391			
Dd	30 ILPLFLYSAGYHASKCEFIANTLHTALTALNVSRRTSEACMTHLHYKWIPIVEDSHADISSH	89			
QY	392 WKETHRFIEAPAAQGTHLVLCMKGVRSRAATVLAYAMKQYESLSLEOALRHVOELRPIAR	451			
Dd	90 FQEAIDFIDCVREKGGKVLVHCAGHSRSPITCMAYLMKTQQRLKEAFDIYIKORSMVS	149			
QY	452 PNPGEFLROLQIYYGILTAS	470			
Dd	150 PNFGFMGOLLQYESELPS	168			

```

RESULT 7
US-09-557-921-19
; Sequence 19, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557,921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-19

```

Query Match	6.4%;	Score 219;	DB 4;	Length 171;
Best Local Similarity	33.8%;	Pred. No. 3.4e-12;		
Matches 47;	Conservative 29;	Mismatches 63;	Indels 0;	Gaps 0

[illegible]

```

RESULT 8
US-09-770-595A-3
; Sequence 3, Application US/09770595A
; Patent No. 6566511
; GENERAL INFORMATION:
; APPLICANT: Revenkova, Ekaterina
; APPLICANT: Paszkowski, Jurek
; TITLE OF INVENTION: Map Kinase Phosphatase Mutant
; FILE REFERENCE: S-30589A
; CURRENT APPLICATION NUMBER: US/09/770,595A
; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-770-595A-3

```

	Query Match	6.4%; Score 218.5; DB 4; Length 784;
	Best Local Similarity	25.5%; Pred. No.4.7e-11;
	Matches	80; Conservative 48; Mismatches 137; Indels 49; Gaps 5
QY	205 QVLHQAACEALGSGLVPGGSALTWASHYOER-----INSEOSCLNEW-TAMADLE 25	
Db	42 QSLNIPCAISSGPSRRCPPAPLTPRHSHNSKARACLPPLQPLAISRSRIDEMPKAGSDDV 10	
QY	254 SLRPPEAEFGSSGEQEOMEQAIRAELWKVLDVSDLES-VTSKEIRQALELRGLPLOOQR 31	
Db	102 GEWPHPPTPSGNKTGERLK-----LDLSSTQQRVTDKS-----SGLAKREKI 14	
QY	313 DFIDNQMLLVAQRDRASRIFFPHLYLGSEMNAANLEELORNRVTHILNMAREI-DNFYPE 37	
Db	144 AFFDKE-----CSKVADHIYVGDAVDAAKDKSILKNNGITHILNCVGFICPEYFKS 19	
QY	372 RFTYHNVRLMDEESAQLLFHWKETHRFTEARAQGTHTVLWHCKMGVSRSATVLAAYAMKO 43	
Db	194 DFCYRSLWLQDSPSEDITSILYDVPDYEDVREQSGRI FVHCQCQGSRSTSIVAIYALMWR 25	
QY	432 YECGLEQALRHVOELRLPIARPMPGFLRQLIQIGILTASRQSHWMEQKGVGVSPEEHPAF 49	
Db	254 EGQSFDDAFOGYVKARSAGTIADPNMGFAQQL-----LQCCQKRVAHFPLSP TSLLRMYKMSP 30	
QY	492 EVSTPFPLPPEPE 505	
Db	308 H--SPYDPHLHVPR 319	

RESULT 9
US-09-704-139-2
; Sequence 2, Application US/09704139
; Patent No. 6420153
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosanna
; TITLE OF INVENTION: 18232, A NOVEL DUAL SPECIFICITY PHOSPHATASE AND USES THEREFOR
; FILE REFERENCE: 10448-018001
; CURRENT APPLICATION NUMBER: US/09/704,139
; CURRENT FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/185,772

PRIOR FILING DATE: 2000-02-29
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 176
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(1390)
OTHER INFORMATION: n = a, t, g, or c
US-09-704-139-2

Query Match 6.3%; Score 215.5; DB 4; Length 176;
Best Local Similarity 38.9%; Pred. No. 7.5e-12;
Matches 61; Conservative 22; Mismatches 71; Indels 3; Gaps 2;

QY 331 RIFPHLYGSEWMAANLELOQRNVTHTILMAEIDNFPYPERFTYHVRMLWDEESAQLLP 390
DB 21 RVAPSLFLGSARAAAGAEQLARAGVTLGVNVSROQPPRAPGVAEIRVPVFDPAEDLLA 80
QY 391 HWKETHRPIEAPAOCTHVLVHCKMGVSRSAATVLAAMKQYECSEQLRHVOELRPIA 450
DB 81 HLEPTCAWEAAVRAAGACLVYCKNGRSRAVCTAYLMRHRGLSLAKAFQWKSARPYA 140
QY 451 RPNPGLRQLQIYOGILTASRQSHV-WEQKGVGSPE 486
DB 141 EPNPGLRQLOKYEALQA--QSCLOGEPALGLGPE 175

RESULT 10
US-08-990-379-5
Sequence 5, Application US/08990379
Patent No. 5998188
GENERAL INFORMATION:
APPLICANT: Stock, Philip J
APPLICANT: Mista-Press, Anita
TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
FILE REFERENCE: 4104-000322USA
CURRENT APPLICATION NUMBER: US/08/990,379
CURRENT FILING DATE: 1997-12-15
EARLIER APPLICATION NUMBER: PCT/US96/10402
EARLIER FILING DATE: 1996-06-14
EARLIER APPLICATION NUMBER: 60/000,263
EARLIER FILING DATE: 1995-06-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 395
TYPE: PRT
ORGANISM: Rattus norvegicus
US-08-990-379-5

Query Match 6.3%; Score 215.5; DB 2; Length 395;
Best Local Similarity 22.7%; Pred. No. 2.8e-11;
Matches 87; Conservative 54; Mismatches 130; Indels 113; Gaps 10;

QY 176 DGDGFSVTSQGSRIFKPISTQTMWATLQVLAQCEAALGSGLVGGSALTW----- 228
DB 22 DENGCTAGSSGGS-----HGAL-----GLISGKCLLDLDRPFLA 56
QY 229 --ASHYQERLNSQSCINEMTA--MADLESIRPPSAEPGSSSEQOMEQAIRAELMWYLD 284
DB 57 HSAGYIRGSVNVRCNTIVRRAKGSVSLQILP-----AEEVVARLRSLGYSAVI 107
QY 285 VSDLESVTSKEIRQALELRGLPLQYRDFID-NQMLLLVAQRDRAS----- 330
DB 108 VYDERSPRAESLRE--DSTVSLVQALRRNAERTDICKGGERFSSEYEPFCSTKRAL 165
QY 331 -----RIFPHLYGSEWMAANLELOQRNVTHT 357
DB 166 AAIIPPVPSTNESLIDGSSCGTPLHDGGPVEILPFLYLGSAYHAARDMDLALGITA 225

QY 358 ILNMAEIDNFPYPERFTYHVRMLWDEESAQLLPHWKETHRFIEAPAOCTHVLVHCKMGV 417
DB 226 LINVSSDCPNHFEGHYQKCIPEVDNHNKADISSWMEAIEXYIDAVDCRGLVHCOAGI 285
QY 418 SRSAATVLAAMKQYECSEQLRHVOELRPIARPNGFLRQ-LQIYOGILTASRQSHV 476
DB 286 SRSATICLAYLMKKRVLEAEAFEFVKQRSIISPNSFMGQLQFESQVLTTS----- 339
QY 477 EQKVGVSPEEHAPAEVSTPPPL 500
DB 340 -----CAAEASPSGPL 351

RESULT 11
US-09-544-716-2
Sequence 2, Application US/09544716
Patent No. 6492157
GENERAL INFORMATION:
APPLICANT: Luche, Ralf M.
APPLICANT: Wei, Bo
TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
FILE REFERENCE: 200125.415
CURRENT APPLICATION NUMBER: US/09/544,716
CURRENT FILING DATE: 2000-04-10
NUMBER OF SEQ ID NOS: 20
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
US-09-544-716-2

Query Match 6.3%; Score 214; DB 4; Length 211;
Best Local Similarity 37.1%; Pred. No. 1.4e-11;
Matches 56; Conservative 26; Mismatches 57; Indels 12; Gaps 4;

QY 327 DRASRIFPHLYGSEWMAANLELOQRNVTHTILMAEIDNFPYPERFTYHVRMLW 381
DB 60 NHADEWPGLYLGDQDMANRRRLRLGITHVLNASHSRWGTPEAYEGLGIRYLGVEAH 119
QY 382 DEESAQLLPHWKETHRPIEAPAO-GTHVLVHCKMGVSRSAATVLAAMKQYECSEQL 440
DB 120 DSPAFDMSIHFOADFIHRLSOPGKILVCAVGSRSATVLAAMLYHMLTLVEAI 179
QY 441 RHVQELRPIARPNGFLRQ-----QIYOGI 466
DB 180 KKVKDHRGII-PNRGFLRQLLALDRRLRGL 209

RESULT 12
US-09-045-973-8
Sequence 8, Application US/09045973
Patent No. 6165767
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

```

1 SOFTWARE: FASTSEQ for Windows Version 2.0
2
3 CURRENT APPLICATION DATA:
4 APPLICATION NUMBER: US/09/045,973
5 FILING DATE: Filed Herewith
6 CLASSIFICATION:
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER:
9
10 FILING DATE:
11
12 ATTORNEY/AGENT INFORMATION:
13 NAME: Billings, Lucy J.
14 REGISTRATION NUMBER: 36,749
15 REFERENCE/DOCKET NUMBER: PF-0491 US
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: (650) 855-0555
18 TELEFAX: (650) 845-4166
19
20 TELEX:
21
22 INFORMATION FOR SEQ ID NO: 8:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 226 amino acids
25 TYPE: amino acid
26 STRANDEDNESS: single
27 TOPOLOGY: linear
28 IMMEDIATE SOURCE:
29 LIBRARY: GenBank
30 CLONE: 1495338
31
32 US-09-045-973-8

```

Query Match	6.3%;	Score 214;	DB 3;	Length 226;
Best Local Similarity	29.7%;	Pred. No. 1.6e-11;		
Matches 69;	Conservative 40;	Mismatches 87;	Indels 36;	Gaps 10;
QY	330	SRI PH YL GS EWNAANLE IQ NRVTHILNMAREIDNFYPERFTYHNVILMDEESAQ--	387	
Db	15	SEIVPG LI CGV-SALSKDEM KK K IK ITHIINATTEVBNL-RSLGDIQ RT KLMLEDTPQTY	72	
QY	388	LLPHWKE TH RFIEARAQ TH VLVHCKMGVSRSAA T VLAYAMKQ Y EC-SIEQALRHVQEL	446	
Db	73	IY PH LEIQSDQIQAL IA DGKVLVHC VA GSRSASIC LA FLK-YRCRNLR EA YHLMKSK	131	
QY	447	RPIARP NG FLRQ LQ IYQ GI L TA SRQSHWEQ K VGV-----SPEE-HPAEVSTPFP	498	
Db	132	RS W VRPNL GF WRQ LI A Y EQNV-----KENAGSV LR VRDEAQPEQLLPDY LI NAIP	182	
QY	499	PLPPEGGGEEKVVGMEESQAAPKEBP RP RI N LRGV MR SISLLEPSLEL	550	
Db	183	ARPASPE-----QDPN MI DE--PRER NS GFKSKFRPVLEPV ME M	221	

```

RESULT 13
US-09-544-716-20
; Sequence 20, Application US/09544716
; Patent No. 6492157
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.415
; CURRENT APPLICATION NUMBER: US/09/544, 716
; CURRENT FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 20
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-716-20

```

	Query Match	6.3%	Score 213.5;	DB 4;	Length 174;
	Best Local Similarity	38.6%;	Pred. No. 1.1e-11;		
	Matches	54;	Conservative	23;	Mismatches 56; Indels 7; Gaps 3.
QY	327 DRASRIFPHYLGSEWNAANLEETQRNRVTHILNMAREIDNYPERF-----TYHNVRLLW	361			

Db 27 NHADVWPGLYLGDDQDMANNRRELRLRIGITHVLNASHSRMRGTPEAYEGLGIRYLGVAAH 86

QY 382 DEESAQLLPWKETHRFIEAARAQ-GTHVLVHCKMGVSRSAATVLAAYAMKÖYECSLEQAL 440

Db 87 DSPAFDMSIHQTADFIHRALSQPGKILVHCavgvsrsatvLayLMLYHHLTLVEAI 146

QY 441 RHVQELRPiARPNPGFLRQL 460

Db 147 KVKXDRGITT-PNRGFLRQL 165

```

RESULT 14
US-09-544-716-14
; Sequence 14, Application US/09544716
; Patent No. 6492157
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-9 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.415
; CURRENT APPLICATION NUMBER: US/09/544,716
; CURRENT FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-544-716-14

```

Query Match	6.2%	Score 211.5	DB 4	Length 170
Best Local Similarity	36.2%	Pred. No. 1.6e-11		
Matches	51	Conservative	24	Mismatches 61
				Indels 5
				Gaps 2
QY	330	SRI FP HYLYLGSENNANLBEELQNRVTHILNMA RE ---DNFYPERFTYHNVR LM DEESA	386	
Db	28	TRILPHLYLGSGQDV LN KDLMTQNGISYVLN AS NCBP DF ICESR FM --RVPINDNYCE	85	
QY	387	QLPHWKETH RF E EA RAQ GT HVLVHCKMGVSRSA TV LAVAMKQYEC S LEQALRHVOEL	446	
Db	86	KLLPWL DK SI EF ID RA KLSSCQVIVHCLAGISRSATIAIAYIMKTGMSSD AY RFYXDR	145	
QY	447	RPIARP NP GF LR Q LT YQ GI L	467	
Db	146	RPSISP NP EN FL GQ LE YERTL	166	

```

RESULT 15
US-09-557-921-15
; Sequence 15, Application US/09557921
; Patent No. 6551810
; GENERAL INFORMATION:
; APPLICANT: Luche, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-10 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.416
; CURRENT APPLICATION NUMBER: US/09/557,921
; CURRENT FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-557-921-15

```

```

Query Match          6.2%; Score 211.5; DB 4; Length 170;
Best Local Similarity 36.2%; Pred. No. 1.6e-11;
Matches 51; Conservative 24; Mismatches 61; Indels 5; Gaps 2;

QY      330 SRIFPHLYLGSEWMNANLEELQRNRVTHIINMAREI--DNFYPERFTYHNVRIMDEESA 386
      :||| ||||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      28  TRILPHLYLGSKDVLNKDMLTQNGISYVLNASNCSCKPDDFICBSRFM--RVPINDNYCE 85

```

Qy	387	QILPHWKETHRFIEAARQCTHYLVHCKMGVRSATVLAAYAMKQYECSL	EQALRHVOEL	446
Db	86	KILPWLDKSI	EFIDRAKILSSCQYIVHCLAGISRSATIAIAYIMKTMGMSDDAYRFVKDR	145
Qy	447	RPIARPMPGFLRQLOIYOGIL		467
Db	146	RPSISPMPN	FLGQLLEYERTL	166

Search completed: January 15, 2004, 06:55:42
Job time : 22 secs